Richard A. Posner

El análisis económico del derecho

NUEVA EDICIÓN
corregida y aumentada

POLÍTICA Y DERECHO

Sección de Obras de Política y Derecho

EL ANÁLISIS ECONÓMICO DEL DERECHO

RICHARD A. POSNER

El análisis económico del derecho

Traducción
EDUARDO L. SUÁREZ

Traducción de la nueva edición
MARÍA TERESA FRANCO GONZÁLEZ
FAUSTO JOSÉ TREJO
DANIEL JUÁREZ HERNÁNDEZ

FONDO DE CULTURA ECONÓMICA

Novena edición en inglés, 2014
Primera edición en español, 1998
Segunda edición en español, 2008
Tercera edición en español a partir
de la novena edición en inglés, 2023

[Primera edición en libro electrónico, 2025]

Posner, Richard A.
El análisis económico del derecho / Richard A. Posner ; trad. de Eduardo L. Suárez, María Teresa Franco González, Fausto José Trejo, Daniel Juárez Hernández — 3ª ed. — México : FCE, 2023
Título original: Economic Analysis of Law
1 200 p. : grafs., cuadros ; 23 × 17 cm — (Colec. Política y Derecho)
ISBN 978-607-16-7962-8

1. Derecho – Economía I. Suárez, Eduardo L., tr. II. Franco González, María Teresa, tr. III. Trejo, Fausto José, tr. IV. Juárez Hernández, Daniel, tr. V. Ser. VI. t.

LC K487.E3 P67 Dewey 340 P656a

Distribución mundial

Esta publicación forma parte del proyecto "Plataformas de difusión científica: narrativas transmedia para México" del Instituto de Investigaciones Dr. José María Luis Mora, proyecto apoyado por el Conahcyt en el año 2023.

CONAHCYT
CONSEJO NACIONAL DE HUMANIDADES CIENCIAS Y TECNOLOGÍAS

D. R. © 1998, Fondo de Cultura Económica
Carretera Picacho-Ajusco, 227; 14110 Ciudad de México
www.fondodeculturaeconomica.com
Comentarios: editorial@fondodeculturaeconomica.com
Tel.: 55-5227-4672

Ésta es una traducción de la novena edición de *Economic Analysis of Law*, de Richard A. Posner, publicada y vendida por el Fondo de Cultura Económica, con el permiso de CCH Incorporated, Nueva York, Estados Unidos, titular de todos los derechos de publicación y venta.

Diseño de portada: Laura Esponda Aguilar

Se prohíbe la reproducción total o parcial de esta obra, sea cual fuere el medio, sin la anuencia por escrito del titular de los derechos.

ISBN 978-607-16-7962-8 (FCE)
ISBN [pendiente] (electrónico-pdf, FCE)
ISBN [pendiente] (electrónico-pdf, Conahcyt)

Impreso en México • *Printed in Mexico*

SUMARIO

Prefacio 9

Primera Parte
El derecho y la economía: introducción

I. *La naturaleza del razonamiento económico* 15
II. *El enfoque económico del derecho* 45

Segunda Parte
El derecho común

III. *La propiedad* 55
IV. *Derechos y remedios contractuales* 120
V. *El derecho familiar y el derecho sexual* 194
VI. *El derecho aplicable a los cuasidelitos* 229
VII. *El derecho penal* 301
VIII. *El derecho común, la historia del derecho y la jurisprudencia* . . 351

Tercera Parte
Regulación pública del mercado

IX. *La teoría del monopolio* 391
X. *Las leyes antimonopólicas* 410
XI. *La propiedad intelectual* 468
XII. *La regulación de la relación del empleo* 497
XIII. *El servicio público y la regulación del medio de transporte colectivo* 537
XIV. *La elección entre la regulación y el derecho común* 568

Cuarta Parte
El derecho de las organizaciones mercantiles y los mercados financieros

XV. *Corporaciones, financiamiento asegurado y no asegurado, quiebra* 615
XVI. *Los mercados financieros* 684

Quinta Parte

EL DERECHO Y LA DISTRIBUCIÓN DEL INGRESO Y LA RIQUEZA

XVII. *Las desigualdades del ingreso, la justicia distributiva y la pobreza* 737
XVIII. *La tributación* 769
XIX. *La transmisión de la riqueza tras la muerte* 816

Sexta Parte

EL PROCESO LEGAL

XX. *El mercado, el sistema acusatorio y el proceso legislativo como métodos de la asignación de recursos* 839
XXI. *El proceso de elaboración de las normas jurídicas* 882
XXII. *Procedimiento civil y penal* 897
XXIII. *La prueba* 971
XXIV. *La aplicación de la ley y el proceso administrativo* 996

Séptima Parte

LA CONSTITUCIÓN Y EL SISTEMA FEDERAL

XXV. *La naturaleza y las funciones de la Constitución* 1033
XXVI. *El debido proceso económico* 1047
XXVII. *La economía del federalismo* 1057
XXVIII. *La discriminación racial* 1086
XXIX. *La protección de los mercados libres en las ideas y en la religión* 1103
XXX. *Cateos, confiscaciones e interrogatorios* 1128

Octava Parte

DERECHO COMPARADO E INTERNACIONAL

XXXI. *Derecho comparado e internacional* 1141

Índice analítico 1161
Índice general 1187

PREFACIO

Ésta es la novena edición de un libro de texto-tratado que se publicó por primera vez en 1973. Su tema es el análisis económico de las reglas y las instituciones jurídicas, o, como se le llama más comúnmente, "el derecho y la economía". Éste es el campo interdisciplinario más destacado de los estudios jurídicos. De su importancia para la comprensión moderna del derecho dan cuenta el número de revistas dedicadas al tema,[1] la cantidad de información enciclopédica al respecto[2] y el número de libros y series de libros (con frecuencia artículos vueltos a publicar) que lo cubren.[3] Además, en las revistas tradicionales de derecho editadas por estudiantes aparecen un gran número de artículos que aplican la economía al derecho.[4]

El análisis económico del derecho ha influido en el derecho en muchos ámbitos, y también ha transformado la manera en que se enseña. La razón tiene que ver con la imagen extremadamente complicada que a primera vista

[1] *Journal of Law and Economics*, *Journal of Legal Studies*, *Journal of Law, Economics and Organization*, *American Law and Economics Review*, *Supreme Court Economic Review*, *European Journal of Law and Economics*, *International Review of Law and Economics*, *Review of Law and Economics*, *Research in Law and Economics, Asian Journal of Law and Economics*, *Journal of Advanced Research in Law and Economics*, *Law and Economics Yearly Review*, *Journal of Competition Law and Economics*. La revista editada por la facultad de derecho de la Escuela de Derecho de Harvard, el *Journal of Legal Analysis*, se centra en el análisis económico del derecho, a decir de sus primeros cuatro volúmenes.

[2] *The New Palgrave Dictionary of Economics and the Law* (Peter Newman [comp.], 1998) (3 vols.) y la *Encyclopedia of Law and Economics* (Boudewijn Bouckaert y Gerrit De Geest [comps.], 2ª ed., 2009) (12 vols.).

[3] Véanse entre otros Steven Shavell, *Foundations of Economic Analysis of Law* (2004); *Research in Law and Economics* (James Langenfeld [comp.], 2012) (25 vols.); *Research Handbooks in Law and Economics* (Richard A. Posner y Francesco Parisi [comps.]) (8 vols.); *Economic Approaches to Law* (Richard A. Posner y Francesco Parisi [comps.]) (30 vols.); *Economics of Legal Relationships* (Nicholas Mercuro y Michael D. Kaplowitz [comps.]) (14 vols.); *Handbook of Law and Economics* (A. Mitchell Polinsky y Steven Shavell [comps.], 2007) (2 vols.).

[4] Algunos ejemplos recientes son David S. Abrams, "The Imprisoner's Dilemma: A Cost-Benefit Approach to Incarceration", 98 *Iowa L. Rev.*, 905 (2013); Jennifer H. Arlen, "Contracting over Liability: Medical Malpractice and the Cost of Choice", 158 *U. Pa. L. Rev.*, 957 (2010); Omri Ben-Shahar, "Fixing Unfair Contracts", 63 *Stan. L. Rev.*, 869 (2011); Anthony Casey, "The Creditors' Bargain and Option-Preservation Priority in Chapter 11", 78 *U. Chi. L. Rev.*, 759 (2011); Keith N. Hylton, "An Economic Perspective on Preemption", 53 *Boston College L. Rev.*, 203 (2012); Daniel M. Klerman, "Personal Jurisdiction and Product Liability", 85 *S. Cal. L. Rev.*, 1551 (2012); Robert H. Sitkoff, "The Economic Structure of Fiduciary Law", 91 *B. U. L. Rev.*, 1039 (2011); Todd J. Zywicki, "The Economics and Regulation of Bank Overdraft Protection", 69 *Wash. And Lee L. Rev.*, 1141 (2012).

presenta el derecho. El derecho se divide en muchas áreas, cada una con su propia y compleja estructura de reglas. Tradicionalmente, las áreas se estudian de manera aislada unas de otras y dentro de cada área las reglas también tienden a estudiarse como corrientes de pensamiento separadas, a menudo encerradas dentro de sus cotos. Sin embargo, un discreto puñado de doctrinas económicas —como la decisión bajo incertidumbre, los costos de transacción, el análisis costo-beneficio, la aversión al riesgo y las externalidades positivas y negativas— pueden, por su aplicación recurrente a través de las áreas del derecho y las leyes, describir gran parte del régimen jurídico y capacitar al estudiante para que obtenga una comprensión más coherente del sistema: para que capte la relación de sus partes y entienda su unidad esencial, y, una vez logrado esto, atienda de manera competente los nuevos problemas conforme puedan surgir.

Este libro difiere de otros libros sobre el análisis económico del derecho por lo que toca al alcance de su cobertura (*casi* todo el sistema jurídico) y al hincapié que se hace en la regulación de comportamientos fuera del mercado: no sólo ejemplos familiares como los delitos y los accidentes y los juicios, sino también ejemplos menos comunes (para los economistas) como la drogadicción, los robos de obras de arte, los actos sexuales, la maternidad subrogada, los rescates en el mar, la profanación de la bandera, el derecho internacional público, los perdones presidenciales, la teoría democrática, el terrorismo y las observancias religiosas. Difiere también en el principio de la organización, pues se organiza alrededor de los conceptos jurídicos antes que de los económicos. Este enfoque permite que el derecho se vea, se entienda y se estudie como un sistema: un sistema que el análisis económico puede iluminar, revelar como coherente y mejorar. Asimismo, hace posible que la economía se vea como una herramienta para el entendimiento y la reforma de las prácticas sociales, y no sólo como un sistema formal de complejidad matemática atemorizante. De hecho, considero que la matematización de la economía ha ido demasiado lejos. La explicación de la economía que se hace en este libro destaca la unidad, la sencillez y la fuerza de los principios económicos, pero también su sutileza. La presentación es informal, y no supone en el lector ninguna familiaridad previa con la economía o la competencia matemática.

El rápido y continuo crecimiento del análisis económico del derecho desde noviembre de 2010, cuando se envió a la imprenta la octava edición, se refleja en los cambios a lo largo de esta nueva edición, incluidos dos capítulos nuevos: uno sobre la propiedad intelectual y el otro sobre el derecho comparado e internacional; éstas son áreas importantes del derecho cuya cobertura fue insuficiente en ediciones anteriores. Asimismo, en esta edición se presta una mayor atención al comportamiento judicial, ejemplificado de manera realista y explicado en términos económicos. En los otros capítulos hay muchos cambios, tanto de fondo como estilísticos. Se han actualizado

las citas de la literatura de derecho y de economía y —una innovación de envergadura y, espero, bien recibida— se han eliminado las múltiples preguntas intercaladas en el texto de las ediciones anteriores (interrogaciones distintas de los problemas al final de cada capítulo, los cuales se han mantenido). Se me ha convencido de que son una distracción. De modo que, en esta edición, en lugar de las preguntas se presentan las respuestas, lo que debería hacer que el libro sea más claro, más revelador y más accesible para los estudiantes. Como una simplificación adicional, se eliminaron los índices separados para los casos y los autores que se citan en el libro.

En general, se trató de aclarar, reducir y mejorar el libro; de hacerlo menos técnico y más próximo al sentido común, y también de volverlo más ecléctico, menos doctrinario, menos "Escuela de Chicago", más abierto a las críticas del enfoque de la "elección racional" del derecho… críticas que me han convencido, entre otras cosas, de complementar ese enfoque con las penetrantes intuiciones de la psicología.

Por último, deseo agradecer las valiosas contribuciones de mis asistentes de investigación en la Escuela de Derecho de la Universidad de Chicago para esta edición: Andrew Adair, Rajan Aggarwal, Lauren Barnett, Rachel Block, Sam Boyd, Greg Buchak, Sean Cooksey, Kali Frampton, Alan Freedman, Adina Goldstein, Michael Kenstowicz, Xingxing Li, Maria Macia, Adam Solomon, Christopher Tosetti y Michael Zhu.

Richard A. Posner
Noviembre de 2013

Primera Parte

EL DERECHO Y LA ECONOMÍA: INTRODUCCIÓN

I. LA NATURALEZA DEL RAZONAMIENTO ECONÓMICO

Este libro se ha escrito con la convicción de que la economía es una herramienta poderosa para el análisis de un gran conjunto de cuestiones jurídicas, y de que la mayoría de los abogados y los estudiantes de derecho —incluso algunos muy brillantes— tienen dificultad para conectar los principios económicos con problemas jurídicos concretos. Un estudiante toma un curso de teoría de los precios y aprende lo que ocurre con el precio del trigo cuando baja el precio del maíz y con el precio de la tierra de pastos cuando aumenta el precio de la carne, pero no entiende qué tienen que ver estas cosas con el derecho de libre expresión o los accidentes o los delitos o la Regla contra las Perpetuidades o las escrituras de una gran empresa o el matrimonio gay. Este libro se propone fijar la discusión de la teoría económica en cuestiones jurídicas concretas, numerosas y variadas; la discusión de la teoría económica en abstracto se limita a este capítulo.

I.1. Conceptos fundamentales

Muchos abogados creen todavía que la economía es el estudio ominosamente matematizado de la inflación, el desempleo, los ciclos económicos, la política monetaria, el comercio exterior y otros fenómenos macroeconómicos alejados de las preocupaciones diarias del régimen jurídico. En realidad, el campo de la economía es por supuesto mucho más amplio y no sólo estudia el comportamiento de ciertos mercados y las prácticas comerciales. Definida de la manera más amplia, la economía es la ciencia de la elección racional en un mundo —nuestro mundo— donde los recursos son limitados en relación con las necesidades humanas. La tarea de la economía, así definida, consiste en la exploración de las implicaciones de suponer que el ser humano procura en forma racional maximizar sus fines en la vida, sus satisfacciones: dicho sea en términos equivalentes, su "interés propio", lo que no debe confundirse, empero, con el egoísmo. La felicidad (o, para el caso, la miseria) de otra persona puede formar parte de nuestras satisfacciones, aunque a fin de evitar esta confusión los economistas prefieren hablar de "utilidad" antes que de interés propio.

La maximización racional tampoco debe confundirse con el cálculo consciente. La economía no es una teoría de la conciencia. El comportamiento es racional cuando se conforma al modelo de la elección racional, cualquiera que sea el estado mental de quien escoge. Por lo tanto, los animales son ra-

cionales. La racionalidad no significa tampoco omnisciencia. La información es costosa y muchas veces los costos son prohibitivos, en especial cuando la información que le gustaría tener a uno se relaciona con el futuro. La teoría de la elección racional acepta que la capacidad de las personas para procesar la información es limitada: existen costos no sólo para obtener la información, sino también para asimilarla y utilizarla.

Para este libro es fundamental el supuesto adicional de que la persona se esfuerza por aumentar racionalmente al máximo la utilidad en *todas* las áreas de la vida, no sólo en sus asuntos "económicos"; es decir, no sólo cuando se ocupa de la compra y la venta en mercados explícitos. Esta idea —que es esencial para el análisis económico del derecho, porque gran parte de la ley se refiere a comportamientos fuera del mercado (como el delito, el matrimonio y el divorcio, los accidentes y las herencias)— se remonta a Jeremy Bentham en el siglo XVIII y principios del XIX, pero recibió poca atención de los economistas hasta la obra de Gary Becker en los decenios de 1950 y 1960.[1]

Pero un supuesto no debe confundirse con un hallazgo. El supuesto de que las personas son electoras racionales en todas las áreas de la vida es sólo el punto de partida del análisis económico. Sólo se mantiene como una aproximación. Incluso las personas completamente normales no siempre son racionales. No es racional asustarse al ver una película de terror, pero la mayoría de las personas reaccionan así. Y hay situaciones en las que la elección racional queda excluida por la incertidumbre, definida como el riesgo que no se puede calcular.[2] Existe un riesgo calculable de empatar teniendo una flor imperial en un juego de póquer, pero no se puede calcular o incluso estimar la probabilidad de que ocurra un ataque terrorista en los Estados Unidos en los próximos seis meses. La incertidumbre es una característica dominante en la vida humana y la economía no la puede pasar por alto. El matrimonio es un ejemplo. En nuestra sociedad, la probabilidad de que una pareja de casados determinada se divorcie no puede calcularse (en general) en el momento en que se casa.

Las personas toman medidas para enfrentar la incertidumbre; en el caso del matrimonio éstas incluyen la elaboración de acuerdos prenupciales. Si la tasa de divorcio aumenta, entonces, aunque aún es incierta la probabilidad de un divorcio determinado, todavía podemos esperar que dichos acuerdos se vuelvan más comunes. Ésa es una respuesta razonable ante la incertidumbre. Otra de las respuestas es un noviazgo prolongado; se espera que la demora del matrimonio permita a ambas partes obtener información que disi-

[1] Véase Gary S. Becker, *The Economic Approach to Human Behavior*, cap. 1 (1976); Richard A. Posner, *Frontiers of Legal Theory*, 54-57 (2001).

[2] La distinción se introdujo en los libros publicados el mismo año por un distinguido economista estadunidense (el fundador de la Escuela de Economía de Chicago) y el famoso economista inglés John Maynard Keynes. Véase Frank H. Knight, *Risk, Uncertainty, and Profit*, 19-20 y el cap. 7 (1921); Keynes, *A Treatise on Probability*, cap. 3 (1921).

pe o al menos reduzca su incertidumbre respecto a la duración probable del matrimonio entre ellos.

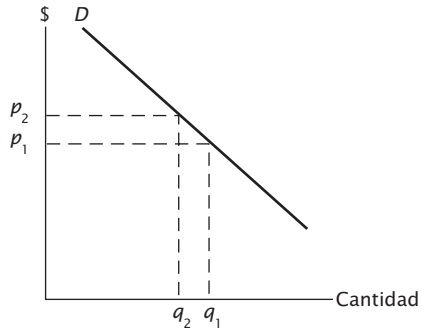
El concepto de incertidumbre aparecerá en varias partes del libro (incluyendo la última sección de este capítulo), pero hagámoslo a un lado por ahora y volvamos al supuesto (muy útil, aunque no exacto) de que los seres humanos son maximizadores racionales de sus satisfacciones. Esto implica que la gente responde a los incentivos; que si cambian las circunstancias de una persona en forma tal que podría aumentar sus satisfacciones alterando su comportamiento, lo hará así. De esta proposición derivan los cuatro principios fundamentales de la economía que organizan el análisis del derecho en este libro. Son la Ley de la Demanda, el costo de oportunidad, la tendencia de los recursos a dirigirse hacia sus usos más valiosos y el equilibrio. Éstos resultan ser premisas claras y a todas luces justificadas del análisis económico.

1. La Ley de la Demanda postula la relación inversa entre el precio cobrado y la cantidad demandada. Si el precio de la carne de res aumenta en 50 centavos por libra, y otros precios permanecen constantes, un bistec le costará ahora más que antes al consumidor, relativamente. El consumidor racional e interesado en sí mismo reaccionará investigando (al menos si ello se puede hacer a un bajo costo en tiempo y esfuerzo) la posibilidad de sustituir el filete con los bienes que prefería menos cuando éste estaba a su antiguo precio pero que ahora resultan más atractivos porque son más baratos en relación con la carne de res. Muchos consumidores continuarán comprando tanta carne de res como antes; para ellos, otros bienes son sustitutos pobres, aunque tengan precios relativos algo menores (o los costos de la búsqueda de sustitutos pueden exceder los beneficios). Pero algunos compradores reducirán sus compras de carne de res y las sustituirán con otros tipos de carnes (u otros alimentos, o productos enteramente diferentes), con el resultado de que la cantidad total demandada por los compradores y, por ende, la cantidad producida disminuirán.

Esto se muestra en la gráfica I.1. En el eje vertical se representan dólares y en el eje horizontal se representan unidades de producción. Un aumento del precio, de p_1 a p_2, se traduce en una disminución de la cantidad demandada, de q_1 a q_2. Igualmente podríamos imaginar que la cantidad ofrecida baja de q_1 a q_2, y observar que el efecto fuese la elevación del precio del bien de p_1 a p_2. La causalidad va en ambas direcciones: si la cantidad disminuye, el precio aumenta; de lo contrario habrá una escasez. Si el precio aumenta, la cantidad disminuye debido a que los consumidores compran una cantidad menor del bien.

Este análisis supone que el único cambio que ocurre en el sistema es el cambio del precio relativo[3] o de la cantidad. Pero si, por ejemplo, la demanda

[3] Si el nivel de los precios está aumentando al mismo ritmo para todos los bienes (es decir, si



GRÁFICA I.1

estuviera aumentando al mismo tiempo que aumenta el precio, es posible que no bajara la cantidad demandada y ofrecida; incluso podría aumentar si el bien se volviera tan popular que lo comprara más gente incluso a su precio más alto.

La Ley de la Demanda no opera sólo en los bienes que tienen precios explícitos. Los maestros impopulares tratan a veces de aumentar el número de sus alumnos elevando la calificación media en sus cursos, con lo que reducen el "precio" del curso para los estudiantes y así los inducen a sustituir otra clase por la clase del maestro impopular. Se dice que el delincuente condenado que ha purgado su sentencia "ha pagado su deuda a la sociedad", y esta metáfora le parecería adecuada a un economista. Desde el punto de vista del delincuente (pero no de la sociedad, a menos que el castigo asuma la forma de una multa), el castigo es el precio que cobra la sociedad por una violación de la ley. El economista pronostica que un aumento en la severidad del castigo o en la probabilidad de su imposición elevará el precio del delito y, por lo tanto, reducirá su incidencia. El delincuente se sentirá estimulado para realizar otra actividad en lugar del delito. Los economistas llaman a los precios no pecuniarios, tales como la desutilidad impuesta por la pena de cárcel, "precios sombra".

2. Supusimos que los consumidores de nuestro ejemplo de la carne de res estaban tratando de maximizar su utilidad (felicidad, placer, satisfacciones).[4] Lo mismo ocurre con los productores de carne de res, aunque en el caso de los vendedores (además de las organizaciones no lucrativas y las dependencias gubernamentales) hablamos de ordinario de la maximización de la ganancia antes que del aumento al máximo de la utilidad. Los vendedores tratan de aumentar al máximo la diferencia existente entre sus costos y sus

hay inflación), no habrá ningún efecto sobre la cantidad, debido a que los precios relativos no cambian, por lo que no existe un incentivo para la sustitución.

[4] En la siguiente sección de este capítulo examinaremos el concepto de utilidad más críticamente.

ingresos por ventas; pero por el momento sólo nos interesa el precio más bajo que cobraría un vendedor racional interesado en sí mismo. Ese mínimo es el precio que exigirían los recursos consumidos en la elaboración (y la venta) del producto del vendedor en su siguiente uso mejor: el precio alternativo. Eso es lo que los economistas entienden por el costo de un bien, y sugiere la razón por la que (sujeto a algunas excepciones que no nos interesan aquí) un vendedor racional no vendería por debajo de su costo. Por ejemplo, el costo de fabricación de una podadora de césped es el precio que el fabricante debe pagar por el capital y la mano de obra consumida en su fabricación. Ese precio debe superar el precio al que podrían haberse vendido los recursos a quien pagase más por ellos en segundo término, porque si el fabricante no hubiese estado dispuesto a dejar atrás ese precio no habría sido el mayor postor y no habría obtenido los recursos. El concepto económico del costo (costo de oportunidad) es el segundo principio fundamental de la economía.

Un corolario de la noción del costo como un precio alternativo es que se incurre en un costo sólo cuando se le niega a alguien el uso de un recurso. Dado que puedo respirar cuanto aire desee sin privar a nadie de nada del aire que desee, nadie me pagará por renunciar a mi aire en su favor, de modo que el aire no tiene ningún costo.[5]

He aquí otros dos ejemplos del costo de oportunidad: *1)* El costo mayor de la educación superior no es la colegiatura, son los ingresos sacrificados que el estudiante tendría si estuviera trabajando en lugar de asistir a la escuela. *2)* Supongamos que los costos de la mano de obra, el capital y los materiales de un barril de petróleo ascienden a sólo 2 dólares, pero, en virtud de que el petróleo de bajo costo se está agotando rápidamente, se espera que la producción de un barril de petróleo cueste 20 dólares dentro de 10 años. El productor que pueda conservar su petróleo durante tanto tiempo podrá venderlo entonces por 20 dólares. Esos 20 dólares son un costo de oportunidad de la venta del petróleo realizada ahora… aunque no un costo de oportunidad neto, pues si el productor espera para vender su petróleo perderá el interés que habría ganado vendiendo ahora e invirtiendo el dinero. Pero supongamos que el precio actual del petróleo es sólo de 4 dólares el barril, de modo que si el productor vende ahora obtendrá una ganancia de sólo 2 dólares. Si invierte los 2 dólares, es improbable que esa suma aumente a 20 dólares (menos el costo de producción entonces) dentro de 10 años. Por lo tanto, el productor se encuentra en mejor situación si deja el petróleo en el subsuelo. Entre más escaso se espere que sea el petróleo en el futuro, mayor será el precio futuro y, por lo tanto, más probable será que el petróleo se deje en el subsuelo, lo que (no es casualidad) tendrá el efecto de impedir una escasez futura.

Esta discusión del costo podría ayudar a despejar una de las falacias más tenaces acerca de la economía: que sólo se ocupa del dinero. Sin duda, el di-

[5] Eso no quiere decir que el aire limpio no tenga ningún costo. Véase después III.7.

nero es una característica esencial de cualquier economía avanzada, como un medio de intercambio que permite la sustitución de un sistema de trueque por un sistema de precios de transacción, como una reserva de valor que se utiliza cuando es necesario para la adquisición de bienes o servicios, y como una medida de valor; pero es sólo una pequeña parte del asunto o tema propio de la economía. Dado que la economía se ocupa principalmente de la asignación de los recursos,[6] los economistas distinguen entre las transacciones que afectan al uso de los recursos, independientemente de que el dinero cambie de manos o no, y las transacciones puramente pecuniarias: los pagos de transferencia, y tienen más que decir acerca de las primeras que sobre estas últimas. Por ejemplo, el trabajo doméstico es una actividad económica, aunque quien trabaje en la casa sea un cónyuge que no recibe ninguna compensación pecuniaria; pues tal actividad implica un costo: primordialmente, el costo de oportunidad del tiempo de quien trabaja en la casa. En cambio, la transferencia por la tributación de 1 000 dólares míos a una persona pobre (o a una persona rica) no es costosa *en sí misma*, es decir, independientemente de sus efectos secundarios sobre sus incentivos y los míos, de los (otros) costos de su ejecución y de cualesquiera diferencias posibles en el valor de un dólar entre nosotros. Esa transferencia no disminuye el acervo de recursos. Disminuye mi poder de compra, pero aumenta en la misma cantidad el poder de compra del receptor. Es lo que los economistas llaman costo *privado*, pero no un costo *social*. Un costo social disminuye la riqueza de la sociedad; un costo privado reordena esa riqueza.

La competencia es una fuente abundante de externalidades "pecuniarias", por oposición a las "tecnológicas"; es decir, de transferencias de riqueza de los individuos sin su consentimiento contrapuestas a la imposición de un costo sobre ellos. Supongamos que A abre una gasolinera frente a la gasolinera de B y en consecuencia atrae ingresos de B. Dado que la pérdida de B es la ganancia de A, no hay ninguna disminución de la riqueza global y, por ende, no hay ningún costo social, aunque B se vea perjudicado por la competencia de A e incurra por lo tanto en un costo privado. Sería diferente si A destruyera las bombas de gasolina de B para eliminar la competencia de éste. Eso reduciría el acervo total de bienes en la economía.

La distinción entre los costos de oportunidad y los pagos de transferencia o, en otras palabras, entre los costos económicos y los contables, nos

[6] Quienes no son economistas otorgan al dinero una importancia mayor que los economistas. Uno de los grandes logros de Adam Smith en *An Inquiry into the Nature and Causes of the Wealth of Nations (La riqueza de las naciones)* fue la demostración de que el mercantilismo, la política de tratar de aumentar al máximo las reservas de oro de un país, empobrecería (en general, aunque no en todos los casos) en lugar de enriquecer al país que la aplicara. Otras concepciones erradas acerca de la economía, que son comunes y que este libro tratará de corregir, son que se ocupa primordialmente de los negocios o los mercados explícitos, que está a favor de los empresarios, que es despiadada, que olvida los costos y beneficios no cuantificables, y que es conservadora.

ayuda a demostrar que el costo es para un economista un concepto orientado hacia delante. Los costos "hundidos" (ya incurridos) no afectan las decisiones de un actor racional sobre el precio y la cantidad. Supongamos que la fabricación de un elefante de porcelana blanco de tamaño natural cuesta 1 000 dólares (éste sería el costo alternativo de los insumos que absorbió su producción), pero, una vez fabricado, nadie pagará más de 10 dólares por el elefante. El hecho de que se hayan "hundido" 1 000 dólares en la fabricación no afectará el precio de su venta, siempre que el vendedor sea racional. Porque si adopta la postura de que no debe vender el elefante por menos de lo que le costó fabricarlo, el único resultado será que en lugar de perder 990 dólares pierda mil.

Esta discusión de los costos hundidos debería ayudarnos a explicar el énfasis que hacen los economistas en la perspectiva *ex ante* (antes del hecho) en vez de en la perspectiva *ex post* (después del hecho). Las personas completamente racionales basan sus decisiones en expectativas del futuro antes que en lamentaciones sobre el pasado. Estas personas tratan el pasado como pasado. Si se permite que las lamentaciones deshagan las decisiones, se verá perjudicada la capacidad de las personas para forjar sus destinos.[7] Si se permite que una de las partes para quien resultó malo un contrato que celebrara libremente revise los términos del contrato *ex post*, se harán pocos contratos.

La aplicación más famosa del concepto de costo de oportunidad en el análisis económico del derecho es el Teorema de Coase.[8] Este teorema, ligeramente simplificado —por lo que toca a las matizaciones necesarias, véase después III.5—, dice que, si las transacciones no cuestan nada, la asignación inicial de un derecho de propiedad no afectará el uso final de la propiedad. Supongamos que un granjero posee el derecho de establecer cultivos que queden librados de los daños causados por las chispas que salen de las locomotoras de un ferrocarril, y que se trata de un derecho de propiedad —es decir, del derecho a excluir la entrada (incluida la entrada de chispas emitidas por las locomotoras) a su tierra, aplicable por un mandato judicial o quizás incluso una sanción penal por la violación de la propiedad— en vez de sólo un derecho a recibir una indemnización por los daños. Supongamos que sus cultivos tienen un valor de 100 dólares y son inflamables, que para el ferrocarril el uso expedito de su derecho de vía tiene un valor de 120 dólares, pero a un costo de 110 dólares puede instalar bloqueadores de chispas (guardachispas) que eliminarán el riesgo de incendios. Bajo estos supuestos, además del supuesto adicional de que a menos que el ferrocarril instale los guardachispas

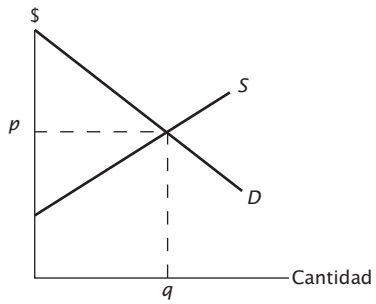
[7] No es la *emoción* del lamento lo que es irracional, sino el hecho de actuar con base en la emoción en lugar de dejar que lo pasado sea pasado. La lamentación es una forma de la autoevaluación que resulta provechosa para mejorar la conducta futura ("no volveré a hacer esto porque ya sé que lo lamentaría"). Véase después I.4.

[8] Ronald H. Coase, "The Problem of Social Cost", 3 *J. Law & Econ.*, 1 (1960).

es un hecho que sus locomotoras destruirán los cultivos del agricultor, el valor real del cultivo para el agricultor no es de 100 dólares, sino alguna suma entre 100 y 110 dólares, porque a cualquier precio mayor a 110 el ferrocarril preferiría instalar guardachispas antes que comprar el derecho de propiedad del agricultor, pero a cualquier precio por debajo de 110 preferiría comprar el derecho de propiedad del agricultor. El agricultor puede obtener el valor más alto del cultivo sólo vendiendo su derecho de propiedad al ferrocarril a un precio entre 100 y 110; hará esto porque el ferrocarril comprará a un precio en ese rango; y, en consecuencia, su tierra será destinada a algún uso que no sea dañado por el fuego, justo como si el ferrocarril hubiese sido su propietario desde un principio. Si no logra vender, sus cultivos serán destruidos, y sólo será compensado por el valor perdido, 100 dólares.

Supongamos ahora que, en vez de que el agricultor tenga el derecho a producir cultivos libres del daño de las chispas de las locomotoras, el ferrocarril tiene el derecho al uso expedito de su derecho de vía. Dado que los cultivos del agricultor tienen un valor menor al costo que implica para el ferrocarril la instalación de los guardachispas, de nuevo las locomotoras arrojarán chispas y la tierra del agricultor se destinará a algún uso que no sea dañado por el fuego. De este modo, la asignación inicial de los derechos de propiedad, ya sea para el ferrocarril o para el agricultor, no afectará el uso final que se haga de la propiedad; en cualquier caso se dispondrá para su uso más productivo, es decir, un uso que no sea dañado por el fuego. La asignación inicial de derechos afectará la distribución de la riqueza, pero no la asignación de recursos. Dicho de otra manera, el costo de oportunidad de los cultivos del agricultor es mayor a su valor; ya que los cultivos son un obstáculo para una actividad económica más valiosa, es decir, la locomoción con emisión de chispas. Para el ferrocarril la tierra vale más si no se cultiva que lo que vale para el agricultor con cultivos en ella.

El proceso por el cual el precio tiende a alinearse con el costo de oportunidad se ilustra en la gráfica I.2. *D* representa la lista de la demanda del bien en cuestión y *S* representa el costo de oportunidad de proveer una unidad del



GRÁFICA I.2

producto a diversos niveles de la producción. Otro nombre para *S* es el de curva del costo marginal de la industria. El costo marginal es el cambio de los costos totales generado por un cambio de la producción en una unidad; en otras palabras, es el costo que se evitará produciendo una unidad menos. (En los capítulos IX y X se explora más a fondo el costo marginal.) Esta definición deberá ayudar a entender por qué la intersección de *D* y *S* es el precio de equilibrio y la producción en equilibrio bajo condiciones de competencia. "Equilibrio" (en pocas palabras) significa un punto estable; es decir, un punto en el que, si no cambian las condiciones de la demanda o de la oferta, no habrá ningún incentivo para que los vendedores alteren el precio o la producción. Si el precio se sitúa a la derecha de este punto, la oferta excederá a la demanda (la curva *S* está por encima de la curva *D* en esta área), de modo que el precio caerá (*S* disminuirá, intersecándose con *D* en un punto más bajo) para permitir que el mercado se despeje (en lugar de acumular los bienes que no se vendieron). A la izquierda del precio de equilibrio y de la producción en equilibrio, donde la demanda es mayor que la oferta, el precio aumentará (*S* aumentará, intersecándose con *D* en un punto más alto, debido a que un precio más alto provoca una mayor oferta) para evitar la escasez que crearía el exceso de demanda si no hubiera un ajuste de precios.

Aun en el equilibrio competitivo a largo plazo no hay seguridad de que se realizarán todas las ventas a precios iguales a los costos de oportunidad de los bienes vendidos. Esto está implícito en la pendiente positiva de la curva de oferta en la gráfica I.2. El hecho de que el costo de producción del bien aumente con la cantidad producida implica que su producción requiere algún recurso que sea escaso en relación con la demanda, como ocurre con la tierra fértil o bien ubicada. Supongamos, por ejemplo, que la mejor tierra de cultivo de maíz puede producir éste a un costo de un dólar por bushel (igual a 25.40 kg), por lo que el costo total consiste a un tiempo en los costos directos de la producción del maíz (mano de obra, fertilizantes, etc.) y en el valor de la tierra en su siguiente mejor uso, y que el precio de mercado del maíz producido en esa tierra sería igual a 10 dólares el bushel si no se produjera más maíz. Es claro que habrá incentivos para expandir la producción, y, dado que la tierra buena no puede expandirse, la tierra inferior se desplazará hacia la producción de maíz: la tierra que requiere mayores insumos de mano de obra, fertilizantes, etc., para la producción de la misma cantidad de maíz.

Este proceso de reasignación continuará hasta que el precio y el costo marginal se igualen, como ocurre en la gráfica I.2. En este punto, el precio de mercado será igual al costo del productor marginal. Supongamos que el costo es 2.50 dólares. Todos los productores de maíz estarán vendiendo el maíz a 2.50 dólares el bushel, pero el costo de la producción de maíz en la mejor tierra será de sólo un dólar. La diferencia existente se llama "renta económica" (no debe confundirse con la renta pagada). ¿Quién se queda con las ren-

tas? No son los productores del maíz, sino los propietarios de la tierra buena (por supuesto, pueden ser la misma persona, pero los papeles de propietario y de productor son distintos). La competencia entre los productores eliminará todas las rentas del productor, dejando toda la renta en manos de los propietarios de los recursos que las generan. Si la cantidad de tierra ideal pudiera aumentarse sin costo, la competencia eliminaría la escasez que genera las rentas, y con ello las rentas mismas. Así pues, bajo la competencia sólo ganan rentas los propietarios de los recursos que no pueden aumentarse rápidamente y a bajo costo para satisfacer una demanda incrementada de los bienes en cuya producción se usan.

Los ingresos muy elevados que ganan unos cuantos cantantes, atletas y abogados incluyen rentas económicas generadas por la escasez inherente de los recursos controlados por estas personas: una excelente voz de cantante, la habilidad y determinación atléticas, las habilidades analíticas y forenses del abogado exitoso. Sus ingresos pueden superar ampliamente a sus mayores ingresos potenciales en una ocupación alternativa, aunque vendan sus servicios en un mercado enteramente competitivo. Una clase diferente de renta económica, discutida en el capítulo IX, es la que gana el monopolista que crea una escasez *artificial* de su producto.

Aunque los economistas tienden a considerar el costo de oportunidad como si fuera lo mismo que el costo de desembolso, existe una diferencia que es aclarada por el concepto de incertidumbre, introducido en este capítulo anteriormente. Un costo de desembolso se observa; un costo de oportunidad es una estimación de lo que ocurrirá si un individuo o una empresa hacen o dejan de hacer algo. Si lo que ocurrirá es incierto, una persona puede dudar en actuar y preferirá el pájaro en mano que el ciento (posible) volando. En los capítulos IX y XV se discute más a fondo la diferencia entre los dos tipos de costos.

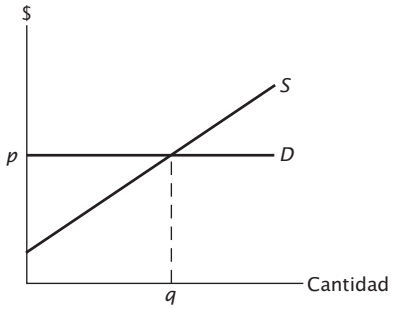
3. Como ya se señaló, el equilibrio es un punto estable. Hasta que las condiciones de la demanda o de la oferta cambian en el mercado, no existe un incentivo para que los vendedores o los compradores abandonen el mercado, o para que entren a él nuevos vendedores o compradores. Pero los equilibrios tienden a ser frágiles y pasajeros. Por ejemplo, imaginemos que el gobierno ha establecido un precio máximo sobre el bien representado en la gráfica I.2 y que el límite máximo está por debajo del precio de equilibrio (de otro modo la regulación sería ineficaz), haciendo así que baje la línea punteada p. En consecuencia, ahora p se intersecará con la curva de la oferta del lado izquierdo de la curva de la demanda, lo que significa que la oferta no satisfará la demanda, debido a que la reducción en el precio disminuye el incentivo de los productores para producir el bien y aumenta al mismo tiempo el deseo de los consumidores de adquirirlo. El resultado es una escasez. El equilibrio se puede restaurar mediante el uso de un método de asignación de la oferta a la

demanda distinto al de los precios. Se podría solicitar a los consumidores que hagan una fila para adquirir el producto; el costo de su tiempo determinaría la longitud de la fila. Las filas son habituales en los mercados en los que se regulan los precios, y en este libro veremos algunos ejemplos. Invariablemente, la eliminación de la regulación de precios reduce, y normalmente elimina, las filas.

La relevancia del concepto de equilibrio es que hace que los economistas piensen bien acerca de las condiciones que impiden que los recursos se dirijan a su uso más valioso y permanezcan ahí de manera indefinida —condiciones que, al introducir una cuña que pende entre el costo y el precio, crean excesos o escasez, y en cualquier caso una pérdida de recursos económicos—. (Por ejemplo, si el precio es mayor al costo, se puede inducir a los consumidores a optar por los productos cuya producción en realidad cuesta más a la sociedad, pero que son más baratos porque se producen en un mercado competitivo, donde el precio iguala los costos y no los sobrepasa.) La percepción de que un mercado en particular no está en equilibrio —por ejemplo, que hay excesos o escasez frecuentes, empresas en liquidación, despidos inesperados y maquinaria que se vuelve obsoleta rápidamente y se desecha, aunque aún funcione bien— incita a la búsqueda de las razones económicas o regulatorias, o incluso psicológicas, por las que el mercado no está en equilibrio y, por consiguiente, no logra asignar los recursos con la máxima eficiencia.

4. Los recursos tienden a gravitar hacia sus usos más valiosos si se permite el intercambio voluntario: un mercado. ¿Por qué el fabricante de podadoras de césped de un ejemplo anterior paga más por la mano de obra y los materiales que los usuarios rivales de estos recursos? La respuesta es que el fabricante pensaba que podría usar tales recursos para obtener un mayor beneficio que el que podrían alcanzar los demandantes rivales. Mediante un proceso de intercambio voluntario, se desplazan los recursos hacia los usos en los que su valor para los consumidores es el más alto, medido por su disposición a pagar. Cuando los recursos se están usando donde su valor es el más alto, o lo que es lo mismo, cuando ninguna reasignación aumentaría su valor, decimos que se están empleando con eficiencia.

Un supuesto metodológicamente útil relacionado con el equilibrio competitivo es que no hay oportunidades de ganancias (en el sentido de las rentas, no del costo del capital social) sin explotar. Una oportunidad de ganancia es un imán que atrae recursos hacia una actividad. Si el imán no funciona, el economista toma esto como una indicación no de que la gente ha dejado de ser racional, sino de que hay algunas barreras para el libre flujo de los recursos. La barrera podría ser la existencia de altos costos de información, las externalidades (efectos, positivos o negativos, que tiene una actividad económica fuera del mercado en el que se lleva a cabo, como cuando los desechos



GRÁFICA I.3

vertidos en un río por una planta que fabrica papel envenenan a los peces), las escaseces inherentes como en nuestro ejemplo de la renta económica de la tierra, u otras condiciones económicas que se discuten en este libro. Si no existen estas barreras, entonces en el mercado representado en la gráfica I.2 cada vendedor confrontará una curva de demanda horizontal (como se muestra en la gráfica I.3) igual a *p*, aunque el mercado en conjunto afronte una curva de demanda de pendiente negativa (la que puede verse como la suma de un número muy grande de curvas de demanda de empresas, cada una de las cuales tiene una pendiente sólo trivialmente negativa, es decir, aproximadamente horizontal, pero cuya suma es marcadamente descendente).[9] La significación de una curva de demanda horizontal es que si el vendedor aumenta su precio, por poco que sea, por encima del precio del mercado, sus ventas bajarán a cero, porque al elevar su precio y abrir así una brecha entre el precio y el costo marginal este vendedor creará una oportunidad de ganancia que otro vendedor le arrebatará.

I.2. VALOR, UTILIDAD, EFICIENCIA

La sección anterior reunió algunas palabras de una alta densidad semántica —*valor, utilidad, eficiencia*— acerca de las cuales debemos ser más precisos.

[9] No es necesario suponer que haya un número infinitamente grande de vendedores en el mercado, sino sólo que la entrada es instantánea si el precio supera al costo marginal. Este punto se explica en el capítulo IX. Adviértase en la gráfica I.3 que la curva del costo marginal de la empresa aparece con una pendiente positiva, justo como la curva de la industria en la gráfica I.2. Las mismas cosas que hacen aumentar el costo marginal de la industria harán aumentar el costo marginal de la empresa; una consideración adicional es el costo creciente de la información y el control a medida que una empresa se vuelve más grande y compleja. Véase después XV.1. Adviértase que si una empresa no encontrara costos marginales crecientes en algún punto, su nivel de producción sería indeterminado. La relación existente entre los costos y la demanda se discute más extensamente en el capítulo XIII.

El *valor* económico de un bien o servicio es lo que alguien esté dispuesto a pagar por ello o, si ya lo tiene, la cantidad de dinero que pide para privarse de ese algo. Éstas no son siempre las mismas cantidades, y esta circunstancia puede causar algunas dificultades, las que consideraremos más adelante. La palabra *utilidad* se emplea con frecuencia en economía para diferenciar un costo o beneficio incierto de uno seguro. La utilidad (más precisamente la "utilidad esperada") en este sentido se liga al concepto de riesgo. Supongamos que le preguntan si preferiría que le dieran un millón de dólares o una oportunidad de 10% de tener 10 millones de dólares. Es probable que prefiriera lo primero, aunque el valor esperado de las dos elecciones sea el mismo: un millón de dólares (= 0.10 × 10 millones de dólares). De ser así, usted siente aversión por el riesgo. La aversión por el riesgo es un corolario del principio de la utilidad marginal decreciente del dinero, lo que significa simplemente que entre más dinero tenga obtendrá menos felicidad adicional con otro dólar. Supongamos que tiene un patrimonio neto de un millón de dólares. ¿Estaría dispuesto a apostarlo con probabilidades de 50-50 para ganar dos millones? Si no está dispuesto, ello significa que su primer millón vale más para usted de lo que valdría un segundo millón.

La aversión al riesgo no es un fenómeno universal; los juegos de azar ilustran su opuesto, la preferencia por el riesgo. Esto no se debe a que uno gane mucho en los juegos de azar sólo por tomar riesgos (ya que son pocas las actividades que realmente no involucran un riesgo), sino a que el valor esperado de un juego de azar suele ser negativo. Supongamos que se le cobra 1 dólar por un billete de lotería, y, dado que se venden un millón de billetes de lotería, la probabilidad de que usted gane la lotería es una en 1 millón. Para que sea un juego justo, el premio por ganar la lotería tendría que ser de 1 millón de dólares, ya que el valor esperado de una en un millón de posibilidades de ganar 1 millón es 1 dólar (= 1 millón de dólares × 0.000001). Pero, de hecho, el premio tendrá que ser menor, pues de lo contrario el organizador de la lotería no tendría ingresos para cubrir sus gastos; recaudaría 1 millón de dólares en la venta de los billetes y pagaría 1 millón de dólares al ganador. Aun así, sabemos que muchas personas compran billetes de lotería. Sin embargo, los economistas creen, con ciertas pruebas —en particular la popularidad de los seguros y lo llamado el "premio de las acciones": el rendimiento esperado mayor de las acciones comunes que de los bonos (véase el capítulo XVI), siendo más riesgosas las acciones ordinarias—, que la mayor parte de la gente siente aversión por el riesgo la mayor parte del tiempo, aunque muchas personas tanto compran seguros como juegan juegos de azar, y aunque veremos que las respuestas institucionales a la aversión por el riesgo, tales como el seguro y la corporación (con su responsabilidad limitada, lo que significa que los accionistas no son responsables de las pérdidas de la corporación), pueden hacer que la gente se vuelva efectivamente neutral al riesgo en muchas situaciones.

Los economistas ya no usan la "utilidad", como lo hicieron alguna vez, en el sentido de que se basa en la teoría filosófica conocida como "utilitarismo", donde la "utilidad" implica el significado aproximado de "felicidad" y se sugiere a la sociedad que trate de maximizar la felicidad promedio o total del ser humano (en algunas versiones la del ser humano más la de los animales). La meta social más destacada en la economía moderna no es la felicidad o la utilidad sino la asignación eficiente de los recursos, en un sentido un tanto especial. Supongamos que A vende a B un adorno de madera en 100 dólares, que ambas partes tienen información plena y que la transacción no afecta a nadie más. Entonces se dice que la asignación de los recursos que se produce por la transacción es superior en el sentido de Pareto a la asignación de recursos existente antes de la transacción. Una transacción superior en el sentido de Pareto (o "mejoramiento de Pareto") es aquella que mejora la situación de una persona por lo menos, y que no empeora la situación de nadie. En nuestro ejemplo, la transacción mejora la situación de A y de B y, según damos por sentado, no empeora la situación de nadie. En otras palabras, el criterio de la superioridad de Pareto es la unanimidad de todas las personas afectadas.

Con todo, las condiciones necesarias para la superioridad en el sentido de Pareto no se satisfacen casi nunca en el mundo real. La mayoría de las transacciones (y si no una sola transacción, entonces una serie de transacciones semejantes) tienen efectos sobre terceros, aunque sólo sea porque cambian los precios de otros bienes (como se señaló en la discusión del concepto de equilibrio) o porque angustian a los observadores (la poligamia, un famoso ejemplo que John Stuart Mill discute en *On Liberty* [*Sobre la libertad*]). En el concepto menos austero de la eficiencia empleado principalmente en este libro —llamado concepto de la eficiencia de Kaldor-Hicks, o de maximización de la riqueza—, si A valúa el adorno de madera en 50 dólares y B lo valúa en 120 dólares, de modo que a cualquier precio entre 50 y 120 dólares la transacción crea un beneficio total de 70 dólares (a un precio de 100 dólares, por ejemplo, A se considera 50 dólares más rico y B se considera 20 dólares más rico), se tratará de una transacción eficiente, siempre que el daño (si hay alguno) causado a terceros (menos cualquier beneficio otorgado a ellos) no exceda de 70 dólares. La transacción no sería superior en el sentido de Pareto a menos que A y B realmente compensaran a los terceros por cualquier daño que sufrieran. El concepto de Kaldor-Hicks se llama también, sugerentemente, la superioridad *potencial* de Pareto: Los ganadores *podrían* compensar a los perdedores, independientemente de que lo hagan o no en realidad.

El hecho de que las condiciones necesarias para la superioridad en el sentido de Pareto no se satisfagan casi nunca en el mundo real y, no obstante, los economistas hablen mucho acerca de la eficiencia significa que la definición operativa de la eficiencia en la economía no puede ser una superioridad

en el sentido de Pareto. Y, de hecho, cuando un economista dice que el libre comercio o la competencia o el control de la contaminación o cualquier otra política o Estado del mundo son eficientes, nueve veces de cada diez se estará refiriendo a la eficiencia de Kaldor-Hicks.

Pero qué pasa si en nuestro ejemplo A valuó el adorno en sólo 70 dólares y B lo valuó en 120 dólares, no porque a A le guste el adorno menos que a B —es posible que le guste mucho más— y no porque haya ningún concepto imperativo de merecimiento al que B pueda apelar para validar su derecho a ser capaz de comprar el adorno, sino porque simplemente A está en la inopia y tiene que vender su adorno a fin de comer, mientras que B, aunque no sea un apasionado de los adornos de madera, sólo desea diversificar su enorme riqueza manteniendo una gama de colecciones. Estas circunstancias no son en modo alguno inconsistentes con el hecho de que la venta mejore la situación de A y de B; por el contrario, explican *por qué* mejora la situación de ambos. Pero tales circunstancias minan los fundamentos morales de un sistema social orientado hacia la superioridad de Pareto, ya no digamos hacia la superioridad potencial de Pareto, al probar que el patrón del consumo y de la producción derivará de una distribución de la riqueza. Si esa distribución es injusta, el patrón de las actividades económicas derivado de ella no tendrá una fuerte justificación para ser considerado como justo tampoco. Y en la medida en que la distribución de la riqueza se determine ella misma por el mercado, la justicia del mercado no puede derivar de alguna noción independiente de la distribución justa.

Además, gran parte de la desigualdad de la riqueza refleja indudablemente la pura suerte: la suerte de haber nacido en un país rico, la suerte de ser un beneficiario o una víctima de los cambios impredecibles que ocurren en los mercados de demandas de los consumidores y de la mano de obra, la suerte de la herencia, la suerte de los mercados financieros, la suerte de conocer a ciertas personas, y la suerte de la capacidad y la disposición de tus padres para invertir en tu capital humano. Entre mayor sea el papel de la suerte en la distribución de la riqueza y las oportunidades económicas, más difícil será defender la distribución como justa en un sentido fuerte, aunque sólo sea en el sentido débil de que no hay una alternativa viable.

Un sistema de mercado moderno tiende efectivamente a *magnificar* las diferencias existentes en materia de la capacidad innata, abriendo una brecha entre la lotería natural y el ingreso. La causa es el fenómeno de la "superestrella".[10] Consideremos a dos pianistas concertistas, uno de los cuales (A) es ligeramente mejor que el otro (B). Supongamos que la mayor parte del ingreso de un pianista concertista no deriva ahora de la actuación o la enseñanza sino de la grabación. Dado que las grabaciones de la misma pieza de música son sustitutos cercanos, un consumidor no tiene ninguna razón

[10] Véase Sherwin Rosen, "The Economics of Superstars", 71 *Am. Econ. Rev.,* 845 (1981).

para comprar las grabaciones hechas por B en lugar de las grabaciones de A a menos que haya una diferencia considerable en el precio, y esto no ocurrirá necesariamente; pues aun si A recibe una regalía mayor que la de B por su contrato con la compañía grabadora, el costo adicional para la compañía grabadora puede ser contrarrestado por las economías de una producción mayor, lo que le permite vender la grabación de A al mismo precio al que esta u otra compañía discográfica podría vender una grabación de la misma pieza musical de B. Por lo tanto, A podría terminar con un ingreso muy sustancial derivado de la grabación y B con un ingreso de cero derivado de la grabación, a pesar de que A podría ser un pianista sólo 2% mejor que B y de que la diferencia de calidad sólo podría ser percibida por un pequeño porcentaje del público melómano.

La definición de *eficiencia* de Kaldor-Hicks genera una paradoja cuando el objeto de una transacción es una gran parte de la riqueza de una persona. Supongamos que rechazo una oferta de 300 000 dólares por mi casa, pero luego el gobierno la confisca, pagándome 200 000, que es su valor de mercado. Con gusto, el gobierno me vendería de vuelta la casa por 250 000 —para el gobierno vale menos que eso, aunque más de 200 000—, pero yo no tengo ni puedo pedir prestados 50 000. ¿En manos de quién vale más la casa, en las mías o en las del gobierno?

Al considerar este enredo, recordemos que la *riqueza*, tal como la entienden los economistas, no es un concepto contable; se mide por lo que la gente *pagaría* por las cosas o por lo que pediría a cambio de renunciar a las cosas que posee, no por lo que *paga* efectivamente por ellas. Regrese a la gráfica I.2 y advierta que si la cantidad vendida fuese menor, el precio sería mayor; evidentemente, los consumidores estarían dispuestos a pagar más por algunas unidades del producto. Por lo tanto, los consumidores deben obtener valor por el hecho de ser capaces de comprar tales unidades al precio competitivo. Este valor, llamado excedente del consumidor (véase después IX.3), forma parte también de la riqueza de la sociedad.

Dado que la economía no responde al interrogante de si la distribución del ingreso y la riqueza existente es buena o mala, justa o injusta, el economista no puede expedir prescripciones obligatorias para el cambio social. Pero esto resulta ser una desventaja menor de lo que podría parecer para el uso normativo del análisis económico. Los economistas pueden apelar de ordinario a una meta generalmente aceptada, como la maximización del valor de la producción, en lugar de tener que defender la meta en cuestión. Al mostrar la forma en que un cambio de una política o de ciertos arreglos económicos nos haría avanzar hacia esa meta, pueden formular una declaración normativa sin tener que defender sus premisas fundamentales. Pueden demostrar, por ejemplo, que la creación de cárteles conduce a una disminución del valor de la producción (véanse los capítulos IX y X), y puesto que la maximización de ese valor es una meta generalmente aceptada de una socie-

dad comercial, su demostración provee, si no otra cosa, una justificación inicial para la prohibición de los cárteles.

Una cuestión normativa importante, a la que ya hemos aludido, es si puede decirse con confianza que un intercambio *involuntario* aumentará la eficiencia y en qué circunstancias. Aun cuando no se defina la eficiencia como algo que sólo puede crear una transacción voluntaria —aun si se utiliza en cambio el concepto de Kaldor-Hicks—, sólo cuando los recursos se desplazan luego de una transacción voluntaria podemos estar razonablemente seguros de que el desplazamiento o viraje implica un aumento de la eficiencia.[11] La transacción no habría ocurrido si ambas partes no hubiesen esperado que su situación mejorara. Pero muchas de las transacciones afectadas o efectuadas por el régimen jurídico son involuntarias. La mayoría de los delitos y de los accidentes son transacciones involuntarias, y lo mismo ocurre con una sentencia de pagar daños o una multa. ¿Cómo podremos saber cuándo aumentan la eficiencia tales transacciones y cuándo la disminuyen? Kaldor-Hicks se pregunta si una transacción voluntaria habría ocurrido en caso de haber sido viable. Si, por ejemplo, se tratara de saber si el agua limpia era más valiosa como un insumo en la producción de papel que en la navegación, podríamos tratar de determinar, empleando cualesquiera datos cuantitativos o de otra clase que estuviesen disponibles para ayudarnos, si en un mundo de cero costos de transacción la industria papelera compraría a los navegantes el derecho de usar el agua.

Es difícil reconstruir los términos probables de una transacción de mercado en circunstancias en las que ocurrió, por el contrario, un intercambio forzado (en otras palabras, un intento de imitar o simular al mercado). Pero la prohibición de todos los intercambios forzados sería muy ineficiente. Esto equivaldría a la abolición de los mercados, ya que prácticamente todas las transacciones de mercado tienen algunos efectos no compensados para terceros.

I.3. El realismo de los supuestos del economista

Ya hemos señalado la falta de realismo de los supuestos clave que subyacen a la teoría económica. El supuesto básico, que el comportamiento humano es racional, parece refutado por las experiencias y las observaciones de la vida diaria, aunque la impresión de falta de realismo es menor cuando uno entiende que el concepto económico de la racionalidad es objetivo, antes que subjetivo; es la capacidad y la inclinación a usar el razonamiento instrumental a fin de desenvolverse en la vida y, por lo tanto, no supone una conciencia (las decisiones racionales suelen tener una justificación patente).

[11] Sin embargo, no podemos estar completamente seguros, porque ello requeriría que todos los afectados por la transacción fueran parte de ella, y (repetimos) este requerimiento no se satisface casi nunca.

Aun con estas reservas, los supuestos de la teoría económica son unidimensionales y deslucidos cuando se contemplan como descripciones del comportamiento humano, especialmente el comportamiento de "actores" económicos tan poco convencionales como el juez, el litigante, el padre, el violador y otros personajes que encontraremos en el análisis económico del derecho. Pero la abstracción es la esencia de la investigación científica, y la economía aspira a ser científica. La ley de Newton de los cuerpos que caen es poco realista en su supuesto básico de que los cuerpos caen en el vacío (así como remplazada por la teoría general de la relatividad), pero sigue siendo una teoría útil porque pronostica con razonable precisión el comportamiento de una gran variedad de cuerpos que caen en el mundo real. De igual modo, una teoría económica del derecho no captará toda la complejidad, la riqueza y la confusión de los fenómenos —criminales o judiciales o maritales o lo que sea— que trata de iluminar. Pero su falta de realismo en el sentido de la totalidad descriptiva, lejos de invalidar la teoría, es una condición previa de la teoría. Una teoría que tratara de reproducir fielmente la complejidad del mundo empírico en sus supuestos no sería una teoría —una explicación— sino una descripción.

Con todo esto no se quiere sugerir que el analista pueda elegir libremente sus supuestos. Una prueba importante de una teoría es su capacidad para explicar la realidad. Si hace un trabajo miserable, ello puede deberse al hecho de que sus supuestos son insuficientemente realistas; pero no necesitamos tratar de evaluar los supuestos directamente a fin de evaluar la teoría. Juzgada por la prueba del poder explicativo, la teoría económica es un éxito significativo (aunque sólo parcial), de modo que quizás el supuesto de que los individuos son maximizadores racionales de sus satisfacciones no es tan poco realista como podría creer a primera vista quien no es economista. La teoría económica puede explicar un gran número de fenómenos del mercado y de fuera del mercado, tales como la correlación inversa, mencionada en la primera sección de este capítulo, entre los topes de precios y las filas; la correlación inversa entre el control de rentas y el acervo de viviendas; la relación existente entre los precios de futuros y los precios de los mercados presentes; la dependencia de la matrícula universitaria respecto de los rendimientos financieros de una educación universitaria; y el hecho de que los mejores bienes tiendan a enviarse a lugares remotos, mientras que los peores bienes se consumen en casa.

Sin embargo, una prueba más de una teoría científica es su poder de predicción, dado que puede ser infinita la cantidad de explicaciones alternativas para el mismo fenómeno. También aquí ha tenido su cuota de éxitos la teoría económica. Por ejemplo, lo ocurrido después del comunismo ha producido repetidas confirmaciones de los pronósticos del análisis económico, como que la regulación de los precios genera la formación de filas, mercados negros y escasez, mientras que la competencia y el libre comercio promueven la pro-

ductividad; y que la propiedad privada estimula la inversión. Y la predicción no tiene que ser de los acontecimientos futuros; puede ser de la forma: La teoría predice que si miramos en un estanque encontraremos una rana. La economía también ha tenido éxitos en este ámbito, así como en la predicción de que los controles de precios generan escasez.

Otra prueba más de una teoría científica es su capacidad para apoyar las intervenciones eficaces en el mundo de la acción. El ejemplo más dramático es el de la bomba atómica, que demostró que la teoría atómica moderna no era simplemente una especulación hábil más acerca de las entidades invisibles. También en este sentido, la economía ha tenido algunos éxitos, aunque más modestos que las ciencias naturales. Los economistas han creado nuevos métodos para la determinación de los precios de productos financieros y de otras clases, nuevas estrategias para la actividad financiera, métodos nuevos para la compensación de empleados y ejecutivos, nuevas técnicas de control de la inflación y el desempleo y para evitar las crisis, así como métodos novedosos de regulación como las subastas de las ondas de radio y los derechos de contaminación vendibles. La mayoría de estas intervenciones han funcionado, aunque no todas… y analizaremos los fracasos, de los cuales el colapso financiero de septiembre de 2008 y la crisis económica mundial resultante se han vuelto muy conscientes los economistas, así como los éxitos.

I.4. Irracionalidad e hiperracionalidad; el origen de la economía del comportamiento y de la teoría de juegos

El hecho de que las personas no siempre tomen decisiones racionales no invalida la teoría de la elección racional; las desviaciones aleatorias del comportamiento racional se cancelarán recíprocamente. Sin embargo, los psicólogos cognitivos, los economistas y el análisis económico del derecho han presentado pruebas de que el comportamiento humano exhibe alejamientos sistemáticos de la realidad.[12] "La percepción y la memoria altamente selectivas dan forma a lo que uno piensa antes de tomar decisiones y de hacer elecciones",[13] e influye así en esas decisiones y elecciones, desviándolas a menudo de lo que predeciría la teoría de la elección racional. Los comunistas exageran la maleabilidad de los seres humanos; los teóricos de la elección racional exageran su racionalidad.

La evolución explica algunas de las deficiencias cognitivas que los teóricos de la elección racional han tendido a ignorar. Por ejemplo, para la mayo-

[12] Véase Daniel Kahneman, *Thinking, Fast and Slow* (2011). Véase en Andrei Shleifer, "Psychologists at the Gate: A Review of Daniel Kahneman's *Thinking, Fast and Slow*", 50 *J. Econ. Lit*., 1 (2012); Ann-Sophie Vandenberghe, "Behavioral Approaches to Contract Law", en *Contract Law and Economics*, 401 (Gerrit De Geest [comp.], 2011) discusiones breves de excelencia.

[13] Shleifer, véase la nota 1 más adelante, en x.

ría de la gente es difícil lidiar racionalmente con sucesos poco probables, y esto bien puede deberse a que en el "ambiente ancestral" (el término utilizado por los biólogos evolutivos para describir el periodo de la prehistoria humana en el que el cerebro humano alcanzó su estado actual) la supervivencia requería atender a tiempo completo las oportunidades y amenazas altamente probables. Otra explicación es que lo que parecen ser deficiencias cognitivas a menudo son atajos cognitivos que funcionan muy bien la mayor parte del tiempo y, por lo tanto, tienen características valiosas de economización. O consideremos la preferencia de un comerciante por fijar el precio de un artículo en $7.99 en lugar de $8. El cerebro humano tiende a fijarse en el primer dígito y, por lo tanto, trata el primer precio como significativamente —antes que trivialmente— menor que el segundo. (Una particularidad cognitiva relacionada explicaría por qué la mayoría de la gente considera que es mucho más fácil comprender las fracciones que los porcentajes, a pesar de que las fracciones y los porcentajes son equivalentes exactos.)

Entre las deficiencias cognitivas que pueden causar estragos en el pensamiento económico racional y, por ende, ocasionar malogros en la elección racional está la falacia de los "costos hundidos" o de soltar dinero bueno tras el malo. Al tomar decisiones ignoramos a menudo el adagio de dejar lo pasado en el pasado; no podemos olvidar los costos que, habiéndose causado ya, no pueden ser modificados por la decisión. Tiene relación con la falacia de los costos hundidos la aversión a la pérdida: con frecuencia se obtiene una utilidad menor de un aumento en el ingreso que la utilidad que se pierde como resultado de la pérdida de una cantidad igual del ingreso. La aversión a la pérdida tiene relación con otra particularidad cognitiva, el "efecto de la dotación": valuar lo que tenemos en más de lo que valuaríamos la misma cosa si nos la ofrecieran. Por ejemplo, podríamos negarnos a vender en 100 dólares un reloj por el que no pagaríamos más de 90 dólares.[14] Adviértase que el efecto de la dotación falsifica al Teorema de Coase, aunque sin afectar su utilidad analítica, que es centrar la atención sobre la importancia de los costos de transacción al explicar el comportamiento de los mercados.

Nos dedicamos a "reflejar", en el supuesto de que otras personas tienen los mismos valores, objetivos y procesos de razonamiento que tenemos, un supuesto que a menudo es falso. También existe el "descuento hiperbólico" (una forma de "falta de voluntad"): ponderamos los dolores y los placeres presentes en mayor medida que los futuros hasta un punto irracional, como cuando comemos de más (placer presente) sabiendo que pronto lo lamentaremos (dolor futuro).[15]

[14] Véase, por ejemplo, Elizabeth Hoffman y Matthew L. Spitzer, "Willingness to Pay *vs.* Willingness to Accept: Legal and Economic Implications", 71 *Wash. U. L. Q.,* 59 (1993).

[15] Más sutilmente, nuestras tasas de descuento son excesivas en relación con nuestro riesgo de mortalidad. Un individuo racional escoge entre acciones posibles usando una tasa de descuento (interés) para disminuir los costos y los beneficios futuros, siempre que se espere que

Otorgamos una ponderación indebida a las impresiones vívidas inmediatas, por comparación con las impresiones acerca de lo que meramente leemos u oímos (la "heurística de la disponibilidad"). Somos demasiado optimistas; en las encuestas, la gente subestima en gran medida la probabilidad de un divorcio, y mucho más del 50% de los conductores consideran que están por encima del promedio respecto a la habilidad de conducción. Estamos excesivamente influenciados por nuestros prejuicios y, por lo tanto, tendemos a ceder ante el "sesgo de confirmación", interpretando pruebas para apoyar una idea preconcebida incluso cuando no lo hacen. Tendemos a la "distorsión retrospectiva", exagerando *ex post* la probabilidad *ex ante* de algún acontecimiento. En retrospectiva tendemos a pensar como inevitable algo que, de hecho, era poco probable que ocurriera; en retrospectiva, no hay accidentes extraños. Una justificación para exigir que se obtenga una orden de registro antes de la búsqueda es, en lugar de dejar que la legalidad de la búsqueda se determine después de que haya ocurrido, que si la búsqueda trae a la luz contrabando o pruebas del delito será difícil para el juez resistir la sensación de que la búsqueda tuvo que haber sido razonable *ex ante*.[16]

En ocasiones, estos alejamientos patentes de la racionalidad pueden explicarse en términos de la elección racional. El optimismo es un antídoto para la depresión debilitante. (Los psiquiatras han acuñado la frase expresiva "realismo depresivo".) También puede ser una condición previa para actuar frente a la incertidumbre, en lugar de experimentar la parálisis.[17] El efecto de la dotación podría ser una consecuencia de alguna combinación de *1)* efectos de riqueza: nuestras preferencias cambian cuando cambia nuestra

lleguen a hacerse realidad, hasta un valor presente, lo que permite una comparación entre los estados futuros y entre cada uno de esos estados y el presente. La imparcialidad entre el consumo presente y el futuro implica el descuento de los costos y los beneficios futuros a una tasa igual a la probabilidad de estar vivo todavía cuando llegue el estado futuro en cuestión. Para la mayoría de los individuos en casi todas las edades, esta probabilidad de sobrevivir es mucho mayor que la que está implícita en una tasa de descuento de 2 a 4%, el rango común de las estimaciones de la tasa de descuento real (esto es, ajustada por la inflación) sin riesgos. El valor presente de un dólar que se recibirá dentro de 40 años es de sólo 21 a 45 centavos a tasas de descuento de 2 a 4%. Esto implicaría, bajo el supuesto de la imparcialidad entre el consumo presente y el futuro, que el individuo típico de 30 años de edad tendría una probabilidad de sólo 21 a 45% de vivir hasta los 70 años. Esta probabilidad es en realidad de 75%. Una posible explicación de esta discrepancia es que el yo de 40 años de edad, orientado hacia el futuro, es incapaz de dominar a su yo orientado hacia el presente, es incompletamente altruista hacia el yo futuro del individuo de 70 años de edad, o ambas cosas. Por lo que toca al individuo como un sitio de diferentes yos, véase más adelante este texto, y Richard A. Posner, "Are We One Self or Multiple Selves? Implications for Law and Public Policy", 3 *Legal Theory*, 23 (1997).

[16] Véase una discusión de este asunto en Richard H. McAdams y Thomas S. Ulen, "Behavioral Criminal Law and Economics" en *Criminal Law and Economics*, 403 (Nuno Garoupa [comp.], 2009).

[17] Véase, por ejemplo, Joel Mokyr, "Intellectual Property Rights, the Industrial Revolution, and the Beginnings of Modern Economic Growth", 99 *Am. Econ. Rev. Papers & Proceedings*, 349, 353 (mayo de 2009).

riqueza, y somos más o menos ricos dependiendo de que seamos o no dueños del bien en cuestión; *2)* el excedente del consumidor (véase antes I.2): una mirada a la gráfica I.2 hará recordar al lector que todos los que son dueños de un bien, excepto su comprador marginal, lo valúan por encima del precio de mercado, por lo que los propietarios del bien como un grupo lo valúan en promedio más que quienes no son sus propietarios; y *3)* la preferencia de adaptación racional: nos adaptamos a lo que tenemos e incurriríamos en costos nuevos para adaptarnos a algo nuevo.[18] Una persona que queda ciega en un accidente debe incurrir en ciertos costos para adaptarse a su ceguera; pero un ciego que por la negligencia de un médico no puede recobrar la vista se ha adaptado ya a su ceguera, de modo que su pérdida de la vista (en perspectiva) es menos costosa que la pérdida de la vista para una persona que no era ciega.

El descuento hiperbólico podría explicarse postulando que el "individuo" es una sucesión de "yos" separados, que "comparten por tiempo" el mismo cuerpo; cada yo es racional, pero cada uno tiene sus propios intereses, los que no son idénticos entre todos los "yos" (véase después VI.12). Incluso la falacia de los "costos hundidos" podría tener una base racional. Sin ella, a menudo la amenaza de tomar represalias no sería un elemento de disuasión creíble (debido a que se ha incurrido en los costos por sufrir un ataque) y, sin embargo, dichas amenazas son una parte importante de la consolidación social. Además, el arrepentimiento, que es la base de la falacia (es difícil "desechar" simplemente las decepciones), nos impulsa a tratar de evitar la repetición de nuestros errores.

Aunque la prueba de que nuestras limitaciones cognitivas inhiben sistemáticamente nuestra capacidad de tomar decisiones racionales es convincente y requiere que se hagan ajustes en el análisis económico tradicional del derecho, existen cuatro razones para moderar el entusiasmo por el "derecho y la economía conductual" (es decir, la inflexión del análisis económico del derecho con el realismo psicológico). La primera es que ignora la autoselección. La mayoría de la gente no sabe tratar con las probabilidades bajas, pero los actuarios de seguros no son un sorteo al azar de la población; se autoseleccionan del subgrupo de personas que saben tratar con probabilidades bajas; de modo que las *compañías* de seguros se comportan de manera racional. Segunda, la experiencia con el mercado o con otras transacciones o actividades tiende a promover el comportamiento racional con respecto a las actividades en las que se involucra la gente con frecuencia.[19] Tercera, las limitaciones cognitivas afectan tanto a jueces, legisladores y otros actores del gobierno como a los particulares… y quizá más, debido a que no tienen tan-

[18] Véase más adelante otra explicación del efecto de la dotación compatible con la racionalidad en III.6.

[19] Véase, por ejemplo, John A. List, "Does Market Experience Eliminate Market Anomaly", 118 *Q. J. Econ.*, 41 (2003).

to en juego, porque no están tan sujetos a los incentivos y limitaciones económicos comunes.[20] Esto limita las consecuencias políticas del derecho y la economía del comportamiento: la capacidad del gobierno para evitar errores atribuibles a las limitaciones cognitivas podría ser bastante limitada. Por último, la aplicación de la economía del comportamiento al análisis económico del derecho puede ser bastante limitada. Tiene aplicaciones importantes en las leyes que regulan las transacciones de los consumidores (véase, por ejemplo, el capítulo XIII de este libro) y, en la versión de esta llamada "finanzas del comportamiento" tiene aplicaciones importantes en la legislación financiera (véase el capítulo XVI). Pero hasta ahora ha tenido una aplicación limitada en la mayoría de los otros campos del derecho.[21]

La teoría de juegos desafía en la dirección opuesta a la teoría de la elección racional. La economía tradicional suponía generalmente (excepto cuando especulaba acerca del comportamiento del cártel, y en otros pocos casos) que los individuos toman decisiones sin considerar las reacciones de otros. Si el precio de algún producto baja, los consumidores compran más sin preocuparse de que al actuar así pueden hacer que el precio aumente de nuevo. La razón por la que no se preocupan es que el efecto de la decisión de cada consumidor sobre el precio es probablemente insignificante (el consumidor es un "tomador de precios"), mientras que los costos para los consumidores de la coordinación de sus acciones serían prohibitivos. En algunas situaciones, sin embargo, un individuo racional a la hora de decidir cómo actuar considerará las reacciones probables de otros; en otras palabras, actuará estratégicamente. Éste es el dominio de la teoría de juegos, la cual exhibe un contraste notable con la economía del comportamiento ya que supone, por lo menos en su forma más pura, un grado de racionalidad mayor incluso que el postulado por la economía ortodoxa.

Consideremos el siguiente "juego" (es decir, una situación estratégica). Un monopolista afronta la perspectiva de la entrada de otra empresa en su mercado. Si el monopolista cobra un precio por debajo de su costo (y presumiblemente el costo de la empresa recién ingresada), impedirá esa entrada obligando al nuevo miembro a perder dinero. Pero el monopolista perderá mucho dinero en el proceso: es posible que pierda más que el que perdería si en vez de bajar el precio de monopolio disminuyera su producción, y se repartiera efectivamente el mercado con el recién ingresado (si el mercado tiene lugar sólo para dos empresas). Se podría pensar que, al saber esto, se incentivaría al neófito potencial para entrar al mercado; pero podría no hacerlo, porque él podría pensar que el monopolista reduciría su precio en lugar de disminuir su producción, pensando lo siguiente: "Si cobro un precio por debajo del costo, el neófito sabrá que no estoy simulando solamente,

[20] Véase Edward L. Glaeser, "Paternalism and Psychology", 73 *U. Chi. L. Rev.*, 133 (2006).

[21] Véase, por ejemplo, Michael G. Faure, *The Impact of Behavioral Law and Economics on Accident Law* (2009).

porque advertirá que aunque perderé dinero a corto plazo, estaré haciendo una inversión rentable en el desarrollo de una reputación que desalentará a otros retadores de mi posición monopólica".

Supongamos ahora que, aunque hay lugar sólo para otra empresa en el mercado, hay diez candidatos que pueden ser esa empresa. Aunque los primeros nueve abandonaran toda idea de ingresar a este mercado frente a la amenaza implícita del monopolista de fijar su precio por debajo del costo, el décimo no lo haría porque advertiría que cuando hay un solo neófito en perspectiva estará el monopolista en una situación mejor compartiendo el mercado que cobrando un precio por debajo del costo. Recordemos que el precio por debajo del costo tenía sentido para el monopolista sólo cuando estaba pensando en comprar una reputación utilizable contra neófitos futuros. Cuando ya no hay más neófitos futuros de que preocuparse, el monopolista ya no tiene nada que ganar invirtiendo en una reputación para bloquear la entrada, de modo que no cobrará un precio por debajo del costo; y, sabiendo esto, el último neófito posible entrará. Pero esto significa que lo mismo hará el primero. Porque sabe que el monopolista, previendo el colapso de su plan cuando entra el último, no tendrá ningún incentivo para emplear el plan contra el noveno (porque, con la entrada del décimo garantizada, el monopolista no tiene nada que ganar forjando una reputación al batir al noveno) y, por lo tanto, contra el octavo, y así sucesivamente hasta llegar al primero. (Esta manera de resolver un problema de teoría de juegos, empezando con la última movida del juego, recibe el nombre de "inducción hacia atrás".)

Lo notable de este ejemplo, algo común en la teoría de juegos, es su sensibilidad ante el supuesto de que cada uno de los participantes no sólo se comporta con una racionalidad completa, sino que además da por sentado, correctamente, que todos los demás involucrados se comportarán igualmente con una racionalidad completa. La teoría económica ortodoxa no depende de tal hiperracionalidad. Pero tampoco lo hacen *todas* las aplicaciones de la teoría de juegos. Consideremos este ejemplo: la ley federal prohíbe que las universidades entreguen a un empleador potencial o a otra institución educativa las calificaciones de un estudiante sin el permiso de éste. Tal permiso no se niega casi nunca. La teoría de juegos puede ayudarnos a entender la razón sin que tengamos que suponer hiperracionalidad. Si ningún estudiante otorgara su permiso, entonces un empleador que afrontara una solicitud de trabajo de un estudiante universitario supondría que tal estudiante tiene calificaciones medias —¿qué otra cosa podría suponer?—. Cualesquiera estudiantes con calificaciones por encima del promedio se verían perjudicados por este supuesto, de modo que empezarían a otorgar su permiso para que las universidades entregaran su registro de calificaciones. A la larga, todos los estudiantes con calificaciones por encima del punto medio otorgarían tal permiso. Por lo tanto, un empleador que recibiera una solicitud de un estudiante que no hubiese entregado sus calificaciones supondría que el estudiante se en-

contraba en el punto medio de la *mitad inferior* de la distribución de calificaciones, porque todos los que se encontraran en la mitad superior habrían revelado sus calificaciones. Así que todo estudiante en el tercer cuartil (es decir, en la mitad superior de la mitad inferior de la distribución de calificaciones) estaría en desventaja al no revelar y no ocultaría sus calificaciones. Finalmente, sólo el estudiante con las calificaciones más bajas no tendría nada que ganar con la revelación, y su incapacidad para dejar la cuestión al descubierto revelaría su lugar tan fehacientemente como si lo hubiera puesto en evidencia. La teoría simple de juegos indica así por qué la ley de privacidad de las calificaciones ha sido ineficaz.

Este ejemplo ilustra lo que los teóricos de juegos llaman un "equilibrio de agrupación", en el que (por oposición a un "equilibrio de separación") el comportamiento estratégico impide que los individuos con preferencias diferentes actúen de manera diferente. El proceso de razonamiento requerido para alcanzar un equilibrio de agrupación en el caso de las calificaciones del estudiante no es tan elaborado que requiera un supuesto de la hiperracionalidad.

I.5. Dos enfoques de la economía

No todos los economistas han estado conformes con la noción de la economía como la ciencia de la elección racional. Los escépticos incluyen a dos de los mejores economistas del siglo XX, Ronald Coase y John Maynard Keynes. Si bien ambos son ingleses, difieren en muchos aspectos, entre ellos que Coase era un microeconomista y Keynes un macroeconomista, y Coase un conservador y Keynes un liberal.

Lo que los une es el escepticismo respecto a la teoría de la elección racional. Coase señaló cosas como la de que "cuando los economistas consideran que no son capaces de analizar lo que sucede en el mundo real, inventan un mundo imaginario que pueden manejar".[22] Escribió que el modelo racional del comportamiento humano es "innecesario y engañoso" debido a que "no existe una razón para suponer que la mayoría de los seres humanos se buscan la maximización de cualquier cosa a menos que se trate de la infelicidad, y aun esto con éxito a medias".[23] Quería que los economistas "estudiaran al hombre tal cual es".[24] Propone abandonar el supuesto de que "las elecciones

[22] Ronald Coase, "The Nature of the Firm: Meaning", 4 *J. Law, Econ. & Organization,* 19, 24 (1988).

[23] Ronald Coase, *The Firm, the Market, and the Law: Essays on the Institutional Structure of Production*, 4 (1988). Se ha referido al modelo racional del comportamiento humano cono "carente de sentido". R. H. Coase, "Coase on Posner on Coase", 149 *J. Institutional & Theoretical Econ.*, 96, 97 (1993).

[24] R. H. Coase, "The New Institutional Economics", 140 *J. Institutional & Theoretical Econ.*, 229, 231 (1984).

de un individuo son coherentes".[25] Rechazó el influyente principio metodológico de Milton Friedman de que una teoría debe probarse no por el realismo de sus supuestos, sino por la exactitud de sus predicciones.[26] Era escéptico acerca de los métodos empíricos, sumamente estadísticos, utilizados en la economía moderna; él prefería el estudio de caso.

Es cierto que Coase estaba asumiendo un comportamiento racional cuando argumentó que la producción se dirige dentro de una empresa y no a través de largos contratos cuando el primer método es más barato, o que si los costos de transacción son nulos la asignación de recursos es independiente de la asignación de los derechos de propiedad a los propietarios. Pero es probable que hubiera dicho que se trataba simplemente de escenarios en los que las personas toman decisiones con base en una comparación entre los beneficios y los costos del plan de acción alternativo. Él no habría estado de acuerdo con los teóricos de la elección racional con respecto a que todas las elecciones humanas tienen este carácter.

Keynes, al igual que Coase, suponía que las empresas comerciales buscan maximizar sus beneficios y los individuos buscan maximizar la utilidad, pero quedó abierto qué tan bien lo lograron. Como Coase, quería ser realista acerca de la toma de decisiones más que explorar hasta qué punto podría llegar un economista al suponer que todas las personas siempre basan sus decisiones en una aproximación cercana al análisis costo-beneficio. Ninguno de los dos se llamaba "científico" a sí mismo.

Una diferencia entre ellos es que mientras que a Keynes le interesaba el comportamiento de los individuos como consumidores y como ahorradores, así como el de las empresas, a Coase no; él se ocupó, casi en su totalidad, de las transacciones comerciales y, por tanto, no tenía necesidad de suponer la maximización de la utilidad: la maximización del beneficio era lo único que necesitaba suponer. Sin embargo, tampoco Keynes utilizaba la maximización de la utilidad excepto en un sentido amplio e implícito. Su obra maestra *The General Theory of Employment, Interest and Money* [*Teoría general de la ocupación, el interés y el dinero*] (1936) está llena de observaciones *psicológicas* interesantes —la palabra "psicológica" es omnipresente—, como cuando señala que "durante el auge, la estimación popular del [riesgo] tiende a volverse inusual e imprudentemente baja", mientras que los "espíritus animales" de los empresarios caen durante el colapso.[27] Keynes confiaba en gran medida

[25] R. H. Coase, Duncan Black, en su *Essays on Economics and Economists*, 185, 190 (1994).

[26] Véase el principio en Milton Friedman, "The Methodology of Positive Economics", en su *Essays in Positive Economics*, 3 (1953). Véase el rechazo de Coase en R. H. Coase, "How Should Economists Choose?" (*American Enterprise Institute*, 1982). Véase también Richard A. Posner, *Overcoming Law*, cap. 21 (1995).

[27] John Maynard Keynes, *The General Theory of Employment, Interest and Money*, p. 130. [Keynes utiliza la figura de "los espíritus animales" para expresar su convicción de que las directrices señaladas por las escuelas clásica y neoclásica, en especial la teoría de Hayek —cuyo funda-

en este tipo de ideas sin preocuparse de por qué tan bien se adaptaban a un modelo de toma de decisiones racionales.

Destacó que a menudo las decisiones de inversión se toman en un escenario de incertidumbre, porque para cuando la inversión comienza a producir un rendimiento las condiciones que determinan su rentabilidad pueden haber cambiado. Aun así, los empresarios hacen inversiones frente a la incertidumbre, y lo hacían antes de que hubiera alguna teoría de la probabilidad. Sin embargo, ¿es "racional" hacer una inversión cuando la propia estimación del beneficio neto esperado es un poco mejor que una puñalada en la oscuridad? La cuestión de la racionalidad no surgió por Keynes. Su tipo de economía no dependía de supuestos muy claros sobre el comportamiento humano. Había observado a los empresarios tomar riesgos no calculables. Si no hubiera personas dispuestas a hacerlo, personas que se sentían "impelidas a la acción", no funcionaría la economía capitalista. El negocio es un campo de actividad atractivo para estas personas; los llaman audaces. Las personas tímidas que son igual de inteligentes que el audaz se vuelven funcionarios públicos, gerentes de nivel medio o profesores.

En un famoso pasaje de la *General Theory*, Keynes había escrito:

> Quizá la mayor parte de nuestras decisiones de hacer algo positivo, cuyas consecuencias completas se irán presentando en muchos días por venir, sólo pueden considerarse como resultado de la fogosidad: de un resorte espontáneo que impulsa a la acción en vez de a la quietud, y no como consecuencia de un promedio ponderado de los beneficios cuantitativos multiplicados por las probabilidades cuantitativas […] De este modo, si la fogosidad se enfría y el optimismo espontáneo vacila, dejando como única base de sustentación la previsión matemática, la empresa se marchita y muere […] Estamos simplemente acordándonos de que las decisiones humanas que afectan el futuro, ya sean personales, políticas o económicas, no pueden depender de la expectativa matemática estricta, desde el momento en que las bases para realizar semejante cálculo no existen; y de que es nuestra inclinación natural a la actividad la que hace girar las ruedas, en tanto nuestro ser racional escoge entre las diversas alternativas lo mejor que puede, calculando cuando hay oportunidad, pero con frecuencia hallando apoyo en el capricho, el conjunto de sentimientos y creencias o el azar.[28]

El debilitamiento de los "espíritus animales" durante una recesión, señalado anteriormente, hace que tanto los negocios como los consumidores re-

mento residía en concepciones presuntamente racionales sobre la cantidad de dinero, el ahorro y la tasa de interés conducentes al "equilibrio"—, conllevaban la ruina de la economía, situación de la que sólo podría salvarse ésta librándose del obstáculo de las deliberaciones eternas y acercándose a los espíritus animales, a saber, ese subsuelo del instinto y el vitalismo en el que se cifran el ánimo resuelto y la capacidad de infundir confianza en las metas por alcanzar. (T.)]

[28] *Ibid*., p. 144.

duzcan el gasto, por temor a lo desconocido. El dinero se atesora, en lugar de gastarse, dando como resultado una caída en el consumo y con él en la producción y el empleo. Un avance importante de Keynes sobre la macroeconomía clásica fue su reconocimiento de que la gente quiere dinero como una reserva ante la incertidumbre, no sólo para el gasto o la inversión; y mientras mayor sea la incertidumbre, más será el dinero que querrán mantener como reserva y, en consecuencia, la economía se volverá más lenta.

El enfoque ecléctico que no establece un límite claro entre la economía y la psicología (y que hoy se vuelve a favorecer con el auge de la economía del comportamiento) llegó a Keynes de manera natural porque era un gran economista, pero no un economista *académico* en el sentido en que se entiende en el siglo XXI. Él no tenía ningún título en economía y combinaba la investigación en Cambridge (no era profesor) con un amplio servicio en el gobierno como asesor y funcionario de alto nivel, y también era un especulador, polemista y periodista.

De este modo, Keynes y Coase fueron metodológicamente similares. Ambos tenían un dominio adecuado, y extendido, de los instrumentos básicos de la economía, pero desdeñaban las matemáticas y, de manera más interesante, el modelo racional del comportamiento humano. Creían que, por lo general, la gente actúa a favor de su interés propio, pero no intentaron hacer una especificación rigurosa de lo que significaba eso. Estaban dispuestos a aceptar a la gente como es en lugar de construir un "hombre racional". No consideraron que el realismo de sus supuestos fuera irrelevante o buscaron comprobar la teoría mediante la exactitud de sus predicciones.

Lecturas recomendadas

Baird, Douglas G., Robert H. Gertner y Randal C. Picker, *Game Theory and the Law* (1994).

Becker, Gary S., "Nobel Lecture: The Economic Way of Looking at Behavior", 101 *J. Pol. Econ.,* 385 (1993).

Behavioral Law and Economics (Jeffrey J. Rachlinski [comp.], 2009).

Coase, Ronald, *The Firm, the Market and the Law: Essays on the Institutional Structure of Production* (1988).

Decision Making: Descriptive, Normative, and Prescriptive Interactions (David E. Bell, Howard Raiffa y Amos Tversky [comps.], 1988).

Demsetz, Harold, "Rationality, Evolution, and Acquisitiveness", 34 *Econ. Inquiry,* 484 (1996).

Friedman, Milton, "The Methodology of Positive Economics", en Friedman, *Essays in Positive Economics,* 3 (1953).

Game Theory and the Law (Eric B. Rasmusen [comp.], 2007).

Harsanyi, John C., "Interpersonal Utility Comparisons", en *The World of Eco-*

nomics: The New Palgrave, 361 (John Eatwell, Murray Milgate y Peter Newman [comps.], 1991).
Hausman, Daniel M., *Essays on Philosophy and Economic Methodology* (1992).
Jolls, Christine, "Behavioral Law and Economics", en *Behavioral Economics and Its Applications*, cap. 3 (Peter Diamond y Hannu Vartieinen [comps.], 2007).
Kaufman, Bruce E., "Some Coasian Problems with Posnerian Law and Economics", 46 *J. Econ. Issues,* 745 (2012).
Keynes, John Maynard, *The General Theory of Employment, Interest and Money* (1936) [edición en español: *Teoría general de la ocupación, el interés y el dinero,* trad. de Eduardo Hornedo, rev. de la trad. de Roberto Ramón Reyes Mazzoni, cuarta edición, Fondo de Cultura Económica, México, 2003; nueva edición revisada próxima a publicarse].
Posner, Richard A., "Normative Law and Economics: From Utilitarianism to Pragmatism", en Posner, *Frontiers of Legal Theory*, 95 (2001).
Shafir, Eldar y Robin A. LeBoeuf, "Rationality", 53 *Ann. Rev. Psych.,* 491 (2002).
Thaler, Richard H., *Quasi Rational Economics* (1991).
The Law and Economics of Irrational Behavior (Francesco Parisi y Vernon L. Smith [comps.], 2005).

Problemas

1. ¿Valdría la pena estudiar economía si los recursos no fuesen escasos? ¿Podría describirse a la sociedad estadunidense actual como una sociedad de escasez? De tiempo en tiempo se afirma que hay escasez de médicos, o de gasolina, o de alguna otra cosa. ¿En qué difieren estos episodios de la escasez de todos los bienes valiosos?

2. Supongamos que todos los individuos fuesen no egoístas, benevolentes y altruistas. ¿Sería entonces el análisis económico menos relevante para el ordenamiento social? ¿Cómo asigna sus recursos el individuo benevolente? ¿Emplearía al mercado —es decir, le vendería a quien pague más— en lugar de regalar sus bienes? ¿Por qué?

3. El mercado es sólo uno de los métodos para destinar la asignación de los recursos a diversos usos. Otro método podría ser el de las decisiones administrativas de un organismo gubernamental. ¿En qué esperaría que difirieran estos métodos?

4. Uno de los costos del castigar el soborno es el de los salarios más altos para los empleados gubernamentales. Explique esto.

5. Explique cómo un bufet de alimentos en un restaurante ilustra el concepto de costo marginal.

6. Se afirma a menudo que la litigación cuesta demasiado y que hay demasiados litigios. ¿Bajo cuáles supuestos pueden ser ciertas ambas proposiciones sin violar la lógica económica?

7. Hay 25 leones (racionales) y un cordero en una isla. Todos los leones conocen el número de leones y el número de corderos, pero no saben dónde se encuentra el cordero. Cualquier león que se coma al cordero se quedará dormido durante una hora, en cuyo lapso puede ser devorado por cualquier león que lo encuentre. ¿Será el primer león que encuentre al cordero el que se lo coma? ¿Sería diferente su respuesta si hubiese 11 leones en lugar de 25? ¿Si hubiera 24 en lugar de 25?

8. Continuando con interrogantes sobre animales, supongamos que le encanta el lechón rostizado. Cierta noche, encuentra un restaurante que sirve el lechón rostizado y lo pide. El mesero le informa que hay varios lechones vivos en la cocina, y querría que escogiera al que le gustaría para que lo mataran y lo rostizaran para usted. Supongamos que el sacrificio y el rostizado se tardan sólo 20 minutos. Rechaza la invitación, sabiendo que si ve vivo al lechón no tendrá corazón para comérselo. ¿Está actuando irracionalmente al estar dispuesto a comerse el cerdo siempre que no tenga que verlo vivo? ¿Sería irracional si, habiendo visto al cerdo, en verdad no pudiera comérselo?

II. EL ENFOQUE ECONÓMICO DEL DERECHO

II.1. Su historia

Hasta finales del decenio de 1960, el análisis económico del derecho era casi sinónimo del análisis económico del derecho antimonopólico, aunque ya se había llevado a cabo en periodos anteriores del siglo algún trabajo económico sobre el derecho tributario (Henry Simons), el derecho corporativo (Henry Manne), el derecho de las patentes (Arnold Plant), el derecho contractual (Robert Hale), y la regulación de los servicios públicos y los transportes colectivos (Ronald Coase y otros).[1] Y en los siglos XVIII y XIX Bentham realizó un trabajo importante sobre la economía del derecho penal,[2] aunque se había olvidado gran parte de él.

Los registros de los juicios antimonopólicos proveyeron una rica mina de información acerca de las prácticas empresariales, y los economistas, que en esa época se preocupaban por la cuestión del monopolio, se dieron a la tarea de descubrir las justificaciones y las consecuencias económicas de tales prácticas. Por supuesto, sus descubrimientos tuvieron implicaciones para el derecho, pero básicamente lo que estaban haciendo no era diferente de lo que habían hecho tradicionalmente los economistas: tratar de explicar el comportamiento de los mercados económicos explícitos.

El análisis económico del derecho antimonopólico, y de otras regulaciones de los mercados económicos explícitos, sigue siendo un campo próspero que recibe una atención considerable en este libro. Sin embargo, la marca distintiva del "nuevo" campo de derecho y economía —el de derecho y economía que surgió en el decenio de 1960 y que aún se mantiene con fuerza— es la aplicación del análisis económico al sistema jurídico en su conjunto: a campos del derecho común, como los cuasidelitos, los contratos y la propie-

[1] Véase más sobre el tema en Ian Ayres, "Discrediting the Free Market", 66 *U. Chi. L. Rev.*, 273 (1999); Charles K. Rowley, "Law and Economics from the Perspective of Economics", en *The New Palgrave Dictionary of Economics and the Law*, vol. 2, pp. 474-476 (Peter Newman [comp.], 1998); Barbara Fried, *The Progressive Assault on Laissez Faire: Robert Hale and the First Law and Economics Movement* (1998); Neil Duxbury, "Robert Hale and the Economy of Legal Force", 53 *Modern L. Rev.*, 421 (1990); Herbert Hovenkamp, "The First Great Law and Economics Movement", 42 *Stan. L. Rev.*, 993 (1990).

[2] Véase Cesare Beccaria, *On Crimes and Punishments* (Henry Paolucci [trad.], 1963); Jeremy Bentham, "An Introduction to the Principles of Morals and Legislation", en *Works of Jeremy Bentham*, vol. 1, pp. 1, 81-154 (John Bowring [comp.], 1843); Bentham, "Principles of Penal Law", en *Works…*, vol. 1, p. 365; Richard A. Posner, "The Law and Economics Movement: From Bentham to Becker", en Posner, *Frontiers of Legal Theory*, cap. 1 (2003).

dad; a campos estatutarios, como la regulación ambiental, la propiedad intelectual, el derecho corporativo y financiero y el derecho de pensiones; a los remedios, como los mandatos y los pagos de daños; a la teoría y la práctica del castigo; al proceso civil, penal y administrativo; a las pruebas; a la teoría de la legislación y la regulación; a la aplicación de la ley y al comportamiento y la administración judicial, e incluso al derecho constitucional, al derecho internacional y comparado, al derecho primitivo, al derecho marítimo, al derecho familiar y a la jurisprudencia.

El nuevo campo de derecho y economía se inició con el primer artículo de Guido Calabresi sobre los cuasidelitos y con el artículo de Ronald Coase sobre el costo social, el cual fue contemporáneo del artículo de Calabresi.[3] Éstos fueron los primeros esfuerzos modernos para aplicar sistemáticamente el análisis económico a áreas del derecho que no regulan las relaciones explícitamente económicas. Podemos encontrar asomos anteriores de un enfoque económico sobre los problemas del derecho de los accidentes y de los perjuicios que examinaron Calabresi y Coase,[4] especialmente en la obra de Pigou,[5] quien proveyó un marco para el análisis de Coase; pero el trabajo inicial había causado escaso impacto sobre el pensamiento jurídico. Comúnmente se había pensado que la economía se define por su objeto más que por su método. Su objeto era la economía; y aunque se sabía que el sistema jurídico afecta a la economía, no se creía que fuera un sistema económico en sí mismo, en el sentido de un sitio de actividad que podría estudiarse en términos económicos de manera provechosa. (El clima afecta a la economía, pero el clima no es un sistema económico.) Como resultado, con la excepción más relevante del análisis utilitarista de la sanción penal de Bentham, había poca literatura económica sobre el derecho fuera de dichas áreas, sobre todo el antimonopolio, en las que el derecho parecía haber incorporado conceptos económicos de manera explícita, tales como el monopolio y la competencia, en el caso del antimonopolio. Un involucramiento mayor, más continuo y más profundo de la economía con el derecho tuvo que esperar a que se reconociera (lo cual había anticipado Bentham nuevamente) que la economía es una teoría, así como un "área de estudio" —en particular la teoría de la elección racional, una teoría que puede aplicarse, en principio, a cualquier actividad social (incluso no humana), incluyendo así al derecho, aun cuando el derecho regula las actividades ubicadas fuera del mercado como el delito, la sentencia o el matrimonio—.

El artículo de Coase introdujo el Teorema de Coase que encontramos en el capítulo I, y, más ampliamente, estableció un marco para el análisis de la

[3] Guido Calabresi, "Some Thoughts on Risk Distribution and the Law of Torts", 70 *Yale L. J.*, 499 (1961); Ronald H. Coase, "The Problem of Social Cost", 3 *J. Law & Econ.*, 1 (1960).

[4] Véanse algunos ejemplos en William M. Landes y Richard A. Posner, *The Economic Structure of Tort Law*, cap. 1 (1987).

[5] A. C. Pigou, *The Economics of Welfare*, cap. 9 (4ª ed., 1932).

asignación de los derechos de propiedad y la responsabilidad en términos económicos, abriendo así un vasto campo de la doctrina jurídica al análisis económico fructífero.[6] Una característica importante del artículo de Coase, aunque pasó inadvertida durante cierto tiempo, era su implicación para el análisis económico positivo de la doctrina jurídica. Coase sugirió que el derecho inglés de los perjuicios tenía una lógica económica implícita. Algunos autores posteriores han generalizado esta idea y sostenido que muchas de las doctrinas e instituciones del sistema jurídico se entienden y se explican mejor como esfuerzos de promoción de la asignación eficiente de los recursos, lo que constituye un tema principal de este libro.

Cualquier lista de los fundadores del "nuevo" campo de derecho y economía estaría incompleta sin el nombre de Gary Becker. La insistencia de Becker en la relevancia de la economía para un conjunto sorprendente de comportamientos fuera del mercado (incluidos la caridad, el amor y la adicción), así como sus contribuciones específicas al análisis económico del delito, la discriminación racial y el matrimonio y el divorcio, abrieron para el análisis económico grandes áreas del sistema jurídico no alcanzadas por los estudios de Calabresi y de Coase sobre los derechos de propiedad y las reglas de la responsabilidad.[7]

II.2. El análisis económico positivo y normativo del derecho

Los capítulos subsecuentes mostrarán la forma en que las penetrantes ideas de los pioneros han sido generalizadas, empíricamente verificadas e integradas a los hallazgos del "antiguo" campo de derecho y economía para crear una teoría económica del derecho amplia dotada de poder explicativo y apoyo empírico. La teoría tiene aspectos normativos y positivos. Aunque el economista no puede decir a la sociedad si debe tratar de limitar el robo, sí puede mostrar que sería ineficiente permitir un robo ilimitado (véase antes I.2); así, puede aclarar un conflicto de valores demostrando cuánto de un

[6] La bibliografía moderna sobre los derechos de propiedad refleja también la influencia de la importante obra anterior de Frank Knight, "Some Fallacies in the Interpretation of Social Cost", 38 *Q. J. Econ.*, 582 (1924); véase después III.1.

[7] Por lo que toca al carácter de las contribuciones de Becker a la economía, véase Gary S. Becker, *The Economic Approach to Human Behavior* (1976); Becker, *A Treatise on the Family* (ed. ampliada, 1991); Becker, *Accounting for Tastes* (1996). Su artículo "Crime and Punishment: An Economic Approach", 76 *J. Pol. Econ.*, 169 (1968) fue de especial importancia para el derecho y la economía. El nuevo campo de derecho y economía es ahora el tema de una bibliografía extensa; además de las porciones pertinentes de este libro, los libros que tratan íntegramente de este tema incluyen los de Steven Shavell, *Foundations of the Economic Analysis of Law* (2004); Robert Cooter y Thomas Ulen, *Law and Economics* (4ª ed., 2003); William M. Landes y Richard A. Posner, *The Economic Structure of Tort Law* (1987); A. Mitchell Polinsky, *An Introduction to Law and Economics* (3ª ed., 2003); Richard A. Posner, *The Economics of Justice* (1981).

valor —la eficiencia— debe sacrificarse para alcanzar otro. O bien, tomando como dada una meta de limitación del robo, el economista puede ser capaz de demostrar que los medios por los cuales ha tratado la sociedad de alcanzar dicha meta son ineficientes: que la sociedad podría obtener mayor prevención a un costo menor empleando métodos diferentes. Si los métodos más eficientes no perjudicaran a otros valores, serían socialmente deseables, aunque la eficiencia ocupara un lugar bajo en el polo totémico de los valores sociales.

Por lo que hace al papel positivo del análisis económico del derecho —el intento de explicar las reglas y los resultados jurídicos tal como son, en lugar de cambiarlos para volverlos mejores—, veremos en los capítulos subsecuentes que muchas áreas del derecho ostentan la marca del razonamiento económico. Si bien pocas opiniones judiciales contienen referencias explícitas a los conceptos económicos, es frecuente que las bases verdaderas de la decisión jurídica estén ocultas antes que iluminadas por la retórica característica de las opiniones que pronuncia la decisión. Se tiene que hurgar debajo de la superficie retórica para encontrar tales bases, muchas de las cuales resultan tener un carácter económico. Esto no es ninguna sorpresa. Los Estados Unidos siempre han sido una sociedad comercial. Muchas doctrinas jurídicas datan del siglo XIX, cuando la corriente del *laissez-faire* basada en la economía clásica era la ideología dominante de las clases educadas. Con la caída del comunismo ha habido un fuerte resurgimiento de la ideología del libre mercado en los Estados Unidos y en gran parte del resto del mundo (el resurgimiento comenzó antes en los Estados Unidos, con el movimiento de desregulación de la década de 1970 y la presidencia de Reagan en la década de 1980), aunque la ideología ha tenido un gran éxito a partir de la crisis económica mundial provocada por el colapso financiero de septiembre de 2008, cuyos efectos aún se dejan sentir.

En los países anglosajones los jueces elaboran la mayoría de las leyes. A esa ley se le llama "derecho común" y veremos que se puede interpretar que gran parte del derecho común tiene como objetivo maximizar la riqueza de la sociedad. El derecho estatutario y el constitucional, por oposición a los campos del derecho común, tienen menores probabilidades de promover la eficiencia, pero aun ellos están impregnados de preocupaciones económicas e iluminados por el análisis económico. Tal análisis es útil también para explicar las características institucionales del sistema jurídico, como el papel del precedente y la repartición de las responsabilidades del cumplimiento forzoso de la ley entre las personas privadas y las dependencias públicas.

Pero podríamos preguntar: ¿no es cierto que el abogado y el economista enfocan el mismo caso en formas tan diferentes que se garantiza una incompatibilidad básica entre el derecho y la economía? X es baleado por un cazador descuidado, Y, y demanda. Lo único que les interesa a las partes y a sus abogados, y lo único que el juez y el jurado decidirán, es si el costo de la le-

sión debiera desplazarse de X a Y; es decir, si es "justo" o "equitativo" que X reciba una compensación. El abogado de X alegará que es justo que X sea compensado, puesto que Y incurrió en falta y X no tuvo ninguna culpa. No sólo son la justicia y la equidad términos no económicos, sino que al economista no le interesa (según creemos) la única cuestión que les interesa a la víctima y a su abogado: ¿quién deberá solventar los costos de *este* accidente? Para el economista, el accidente es un capítulo cerrado. Los costos que infligió son costos hundidos. Al economista le interesa la prevención de los accidentes futuros que no se justifiquen por los costos y, por lo tanto, la reducción de la suma de los costos de los accidentes y de su prevención. Pero las partes del litigio pueden no tener ningún interés por el futuro. Su único interés puede reducirse a las consecuencias financieras de un accidente pasado.

Esta dicotomía es exagerada. La decisión de este caso afectará al futuro, de modo que deberá interesarle al economista porque establecerá o confirmará una regla para la orientación de los individuos que realizan actividades peligrosas. La decisión es una prevención en el sentido de que, si nos comportamos de cierto modo y se produce un accidente, tendremos que pagar una compensación (o no podremos obtener una compensación, si somos la víctima). Al modificar así el precio del comportamiento riesgoso que afronta la gente, la prevención puede afectar su comportamiento y, por ende, los costos de los accidentes.

La decisión del juez será un precedente que influya sobre la decisión de casos futuros. Por lo tanto, el juez también debe considerar el impacto futuro de reglas alternativas. Si el fallo favorece al demandado porque es una persona "meritoria", aunque descuidada, la decisión alentará a individuos similares a ser descuidados. Por lo tanto, una vez que se expande el marco de referencia más allá de las partes inmediatas del caso, la justicia y la equidad asumen significados más amplios de lo que es simplemente justo o equitativo entre este demandante y este demandado. El problema se convierte en la determinación de lo que sea justo y equitativo para una *clase* de actividades, y no puede resolverse sensatamente sin tener en consideración el efecto futuro de reglas alternativas sobre la frecuencia de los accidentes y el costo de las precauciones.

II.3. El enfoque continental

Existe una tradición continental, así como una angloamericana, del pensamiento económico sobre el derecho, la cual surgió en la teoría de la justicia correctiva de Aristóteles (la cual analizamos más adelante en VIII.7), pero se desarrolló plenamente en los escritos de Max Weber y Friedrich Hayek. La primera se diferencia de la tradición angloamericana en cuanto a que hace un mayor énfasis en el derecho como un marco neutral para la conducta privada. Aristóteles hizo énfasis en que un juez no debe tener en cuenta las ca-

racterísticas de los litigantes: los méritos y deméritos, la posición social, las deserciones, etc. Él ha de juzgar el caso, no a las partes. Max Weber insistía en que el propio juez se preocupa de la racionalidad formal: de proveer un marco que facilite el ordenamiento privado en vez de uno que establezca las condiciones en las que las personas lidian entre sí.[8] Es similar la insistencia de Hayek en que los jueces deben obligar al cumplimiento de las costumbres, en otras palabras, las normas creadas por particulares, en lugar de que se establezcan reglas de conducta creadas por los jueces con base en su concepción de una política sólida, es decir, impuesta desde arriba hacia abajo.[9] La tradición continental asigna un papel limitado a los jueces, un papel que se refleja en el énfasis que los sistemas jurídicos continentales asignan a los códigos legislativos como la fuente principal, en ocasiones exclusiva, del derecho, en lugar de la jurisprudencia (es decir, el derecho común). El objetivo es negar a los jueces un papel importante en la formulación de políticas. Los enfoques de Weber y de Aristóteles no están lejos del enfoque angloamericano (de hecho, Weber consideraba que el derecho inglés era superior al continental, al menos en el ámbito comercial). El enfoque de Hayek está lejos de éste.

II.4. Críticas del enfoque económico del derecho

El análisis económico del derecho ha suscitado antagonismo. Una crítica recurrente es que los fundamentos normativos del enfoque económico son tan moralmente repulsivos que resulta inconcebible que el sistema jurídico pueda aceptarlos. ¿Pero es el concepto de eficiencia de Kaldor-Hicks realmente tan diferente de aquellas normas? Es un componente del sistema ético de nuestra sociedad orientada al mercado, y es posible que sea el que domine al derecho tal como lo administran los tribunales, dada la incapacidad de estos últimos para promover eficazmente otras metas.

Otra crítica recurrente del enfoque económico del derecho es que manifiesta un sesgo político conservador.[10] Veremos más adelante que algunos de sus practicantes, en modo alguno todos, han descubierto que la pena capital disuade a alguien, que la legislación diseñada para proteger a los consumidores acaba perjudicándolos con frecuencia, que el seguro del auto sin falta es ineficiente, que cierto nivel de discriminación es eficiente y que buena parte de la regulación de los valores puede ser una pérdida de tiempo. Los libertarios extremos, que quieren liberar al mercado de algunas (en cierta versión

[8] Véase sobre las concepciones aristotélicas y weberianas del derecho Richard A. Posner, *Law, Pragmatism, and Democracy*, 284-287 (2003).

[9] Después, en VIII.4, se analiza el enfoque del derecho de Hayek.

[10] ¡Aunque no lo suficiente para algunos gustos! Véase, por ejemplo, James A. Buchanan, "Good Economics-Bad Law", 60 *Va. L. Rev.,* 483 (1974); Richard A. Epstein, "A Theory of Strict Liability", 2 *J. Legal. Stud.,* 151, 189-204 (1973).

de todas) las regulaciones gubernamentales, tales como David Friedman y Richard Epstein, han ocupado un lugar destacado en el análisis económico del derecho. Sin embargo, un número de analistas económicos del derecho han apoyado muchas posiciones liberales. Por ejemplo, la teoría de los bienes públicos (véase después XVII.4) es uno de los fundamentos ideológicos del Estado benefactor. Varios hallazgos de los analistas económicos del derecho que se discutirán en capítulos subsecuentes de este libro —los relativos al derecho a la asesoría y el estándar probatorio en los juicios penales, a la fianza, a la responsabilidad de los productos, a la aplicación de la Primera Enmienda a la radiodifusión, a los costos sociales del monopolio, a los daños en los casos de lesiones personales, a la regulación del sexo, a las respuestas ante el calentamiento global y a muchos otros— apoyan posiciones liberales. Un considerable número de analistas económicos prominentes del derecho son decididamente liberales, tales como Ian Ayres, Guido Calabresi, John Donohue, Gillian Hadfield y Christine Jolls. Y los analistas económicos conservadores *libertarios* del derecho, como Friedman y Epstein, apoyan posiciones liberales en cuanto al comportamiento personal, como autorizar el matrimonio homosexual y el abandono de la "guerra contra las drogas".

Se critica al enfoque económico del derecho porque se olvida de la "justicia". Al evaluar la crítica debemos distinguir entre los diferentes significados de esta palabra. A veces significa la justicia distributiva (el grado apropiado de igualdad económica) y a veces significa la eficiencia. Veremos más adelante, entre otros ejemplos, que cuando la gente describe como injusta la condena de una persona sin un juicio, la confiscación de la propiedad sin una compensación justa, o el olvidarse de hacer que el automovilista negligente pague los daños a la víctima de su negligencia, no se refiere a algo de mayor alcance que el argumento de que esa conducta desperdicia recursos (véase más adelante VIII.8). Aun el principio del enriquecimiento injusto puede derivarse del concepto de la eficiencia (véase después IV.14). No resultará sorprendente que en un mundo de recursos escasos se considere inmoral el desperdicio.

Pero las nociones de justicia *son* algo *más* que una preocupación por la eficiencia. No es obviamente ineficiente que se permitan los pactos suicidas; que se permita la discriminación privada por razones raciales, religiosas o sexuales; que se permita matar y comerse al pasajero más débil del salvavidas en circunstancias de desesperación genuina; que se obligue a los individuos a declarar en contra de sí mismos; que se azote a los prisioneros; que se permita la venta de bebés para adopción; que se permita la tortura para obtener información; que se permita el uso de fuerza mortal en defensa de un interés puramente propietario; que se legalice el chantaje; que se permita la poligamia o el matrimonio homosexual, o que se haga elegir a los delincuentes condenados entre la prisión y la participación en peligrosos experimentos médicos. Aun así, todas estas cosas ofenden el sentido de justicia de muchos, en ocasiones de la mayoría o incluso de todos, de los estadunidenses

modernos, y todas son ilegales en mayor o menor (de ordinario mayor) medida. En este libro trataremos de explicar algunas de estas reacciones en términos económicos, pero la mayoría de ellas no pueden serlo; evidentemente, la justicia es algo más que economía, lo que el lector deberá tener presente al evaluar los juicios normativos de este libro.

Lecturas recomendadas

Chicago Lectures in Law and Economics (Eric A. Posner [comp.], 2000).
Coase, Ronald H., "The Problem of Social Cost", 3 *J. Law & Econ.,* 1 (1960).
"Ethics, Economics, and the Law", 24 *Nomos* (1982).
Friedman, David D., *Law's Order: What Economics Has to Do with Law and Why It Matters* (2001).
Kennedy, Duncan, "Cost-Benefit Analysis of Entitlement Problems: A Critique", 33, *Stan. L. Rev.,* 387 (1981).
Leff, Arthur Allen, "Economic Analysis of Law: Some Realism About Nominalism", 60 *Va. L. Rev.* 451 (1974).
Mercuro, Nicholas, y Steven G. Medema, *Economics and the Law: From Posner to Post-Modernism and Beyond* (2ª ed., 2006).
Miceli, Thomas J., *Economics of the Law: Torts, Contracts, Property, Litigation* (1997).
Posner, Richard A., "The Law and Economics Movement: From Bentham to Becker", en Posner, *Frontiers of Legal Theory*, cap. 1 (2001).
"Symposium: Law and Economics and Legal Scholarship", 79 *Chi.-Kent L. Rev.* 345 (2004).
The Origins of Law and Economics: Essays by the Founding Fathers (Francesco Parisi y Charles K. Rowley [comps.], 2005).

Segunda Parte

EL DERECHO COMÚN

III. LA PROPIEDAD[1]

Este capítulo inicia nuestro examen del derecho común, un término ambiguo, como muchos otros términos jurídicos. Se emplea de diversas maneras para hacer referencia: *1)* al cuerpo de principios aplicados por los tribunales reales de Inglaterra (excluyendo así al derecho de la equidad y al derecho marítimo, pero incluyendo a gran parte del derecho estatutario); *2)* a los campos del derecho que han sido creados en gran medida por los jueces como un subproducto de la decisión de los juicios, antes que por las legislaturas (incluyendo así al derecho de la equidad y al derecho marítimo), y *3)* a cualquier campo del derecho forjado en gran medida por los precedentes judiciales (incluyendo así a la mayor parte del derecho antimonopólico y grandes muestras del derecho de la propiedad intelectual, grandes áreas de otras esferas jurídicas, e incluso gran parte del derecho constitucional). Esta parte del libro se ocupa del derecho común principalmente en su segundo sentido, pero excluye dos áreas importantes de la legislación elaborada (en gran parte) por jueces: las leyes procesales y el conflicto de leyes, que se discuten en el capítulo xxii.

El derecho común en el segundo sentido, cuando se contempla desde el punto de vista económico, tiene cinco partes:

1. El derecho de la propiedad (con excepción de la propiedad intelectual, que es estatutaria en su mayoría, aunque no por completo), que se ocupa de la creación y la definición de los derechos de propiedad, que son derechos para el uso exclusivo de recursos valiosos;
2. El derecho contractual, que se ocupa de la facilitación del movimiento voluntario de los derechos de propiedad hacia aquellos que los valúan más;
3. El derecho aplicable a los cuasidelitos, que se ocupa de la protección de los derechos de propiedad en un sentido amplio, incluido el derecho a la integridad corporal;
4. El derecho penal: principalmente estatutario, pero la mayoría de cuyos conceptos básicos (como el intento y la conspiración) pertenecen al derecho común; desde un punto de vista económico, el derecho penal es en gran parte, como veremos en el capítulo vii, un apoyo para el derecho aplicable a los cuasidelitos;

[1] John E. Cribbet y otros, *Property Cases and Materials* (9ª ed., 2008); Robert C. Ellickson, Carol M. Rose y Bruce A. Ackerman, *Perspectives on Property Law* (3ª ed., 2002); Sheldon F. Kurtz y Herbert Hovenkamp, *Cases and Materials on American Property Law* (4ª ed., 2003).

5. Doctrinas correctivas y complementarias, como el pago de daños, la restitución, los mandatos judiciales y la sanción penal.

III.1. La teoría económica de los derechos de propiedad: aspectos estáticos y dinámicos

Para entender la economía de los derechos de propiedad es necesario comprender primero la distinción entre el análisis *estático* y el análisis *dinámico*. El análisis estático suprime la dimensión temporal de la actividad económica: se supone que todos los ajustes al cambio ocurren simultáneamente. Este supuesto es poco realista, pero a menudo resulta fructífero.

El análisis dinámico, en el que se relaja el supuesto del ajuste instantáneo al cambio, es más complejo que el análisis estático. Es por ello sorprendente que la base económica de los derechos de propiedad se percibiera al principio en términos dinámicos. Imaginemos una sociedad en la que han sido abolidos los derechos de propiedad. Un agricultor siembra maíz, lo fertiliza e instala espantapájaros; pero cuando el maíz está maduro, su vecino lo cosecha y se lo lleva para su propio uso. El agricultor no tiene ningún recurso legal contra la conducta de su vecino porque, aunque la propiedad implica el derecho a excluir a otros de tomar o usar una propiedad sin consentimiento de su propietario, no es propietario de la tierra que sembró ni del cultivo. A menos que sean viables algunas medidas defensivas, después de unos cuantos incidentes como éste se abandonará el cultivo de las tierras y la sociedad empleará métodos de subsistencia (como la caza) que implican menos inversión preparatoria.

Aunque es posible que el valor del cultivo en nuestro ejemplo, medido por la disposición de los consumidores a pagar, haya excedido ampliamente su costo en términos de la mano de obra, los materiales y los usos alternativos sacrificados de la tierra, sin derechos de propiedad no hay ningún incentivo para incurrir en estos costos porque no hay ninguna recompensa razonablemente asegurada. Los incentivos apropiados se crean parcelando derechos mutuamente excluyentes para el uso de recursos particulares entre los miembros de la sociedad. Si cada parcela de tierra es propiedad de alguien —si hay siempre alguien que puede excluir a todos los demás del acceso a cualquier área dada—, los individuos tratarán de maximizar su valor mediante el cultivo u otros mejoramientos. La tierra es sólo un ejemplo. El principio se aplica a todos los recursos valiosos.

Todo esto ha sido bien sabido durante siglos.[2] En cambio, el análisis es-

[2] Véase, por ejemplo, 2 William Blackstone, *Commentaries on the Laws of England,* 4, 7 (1766). Además, los sistemas de derechos de propiedad tienen orígenes prehistóricos; Vernon L. Smith, "The Primitive Hunter Culture, Pleistocene Extinction, and the Rise of Agriculture", 83 *J. Pol. Econ.,* 727 (1975).

tático de los derechos de propiedad tiene sólo 90 años de antigüedad.[3] Imaginemos que varios agricultores son propietarios de un pastizal en común; es decir, ninguno tiene el derecho de excluir a los otros y, por ende, ninguno puede cobrar a los demás por el uso de los pastos. Podemos olvidarnos de los aspectos dinámicos del problema suponiendo que el pasto es natural (no cultivado), de modo que no hay manera de mejorarlo mediante la inversión. Aun así, el pastoreo de vacas adicionales impondrá un costo a todos los agricultores. Las vacas tendrán que caminar más a fin de comer la misma cantidad de zacate, y esto reducirá su peso. Pero en virtud de que ninguno de los agricultores paga por usar el pasto, ninguno tomará en cuenta este costo al decidir cuántas vacas adicionales pastarán, de modo que pastarán más vacas de lo que sería eficiente. (Adviértase la analogía con el congestionamiento carretero.)

El problema desaparecería si una persona fuera propietaria del pasto y cobrara a cada agricultor por su uso. El cobro a cada agricultor incluiría el costo que éste impone a los otros agricultores al pastorear vacas adicionales, porque ese costo reduce el valor del pasto para los demás agricultores y, por lo tanto, el precio que están dispuestos a pagar al propietario por el derecho de pastar. (Adviértase la analogía con la autopista de peaje.)

La creación de derechos de propiedad individuales, en oposición a los colectivos, es una condición necesaria pero no suficiente para el uso eficiente de los recursos. Supongamos que el agricultor de nuestro primer ejemplo es propietario de la tierra que siembra, pero es un mal agricultor; su tierra sería más productiva en las manos de otro. La eficiencia requiere un mecanismo por el que pueda inducirse al agricultor a transferir la propiedad a alguien que pueda trabajarla más productivamente. Un derecho de propiedad transferible es tal mecanismo. Si la tierra tiene un valor de 100 000 dólares para su propietario actual y de 200 000 para otra persona, entonces la venta de la tierra a cualquier precio intermedio hará que ambas personas mejoren su situación.

Podría parecer que la discusión de este punto implica que se maximiza el valor si todos los recursos valiosos (es decir, escasos y deseados) son propiedad de alguien (el criterio de la universalidad), si la propiedad connota el poder irrestricto para excluir a todos los demás del uso del recurso (la exclu-

La proposición de que la aplicación de los derechos de propiedad conducirá a una producción mayor es cuestionada por Frank I. Michelman en "Ethics, Economics, and the Law of Property", 24 *Nomos,* 3, 25 (1982). Sugiere Michelman que el agricultor que sabe que le robarán la mitad de su cosecha simplemente sembrará el doble. Esta sugerencia olvida: *1)* el incentivo adicional para el robo que se creará al sembrar más y la probabilidad resultante de que se roben más de la mitad del cultivo incrementado; *2)* la poca probabilidad de que la agricultura fuese mucho más rentable que otras actividades sustitutas que no impliquen una inversión preparatoria de modo que la gente siga dedicándose a la agricultura ante el espectro del robo rampante, y *3)* la probabilidad de que el agricultor, en el caso de que permanezca en la agricultura, desvíe algunos de los recursos del cultivo a su protección con vallas, guardias, etcétera.

[3] Véase Frank H. Knight, "Some Fallacies in the Interpretation of Social Cost", 38 *Q. J. Econ.,* 582 (1924).

sividad), así como para usarla uno mismo, y si los derechos de propiedad son libremente transferibles, o alienables, como dicen los abogados (la transferibilidad). Pero esto no toma en cuenta los costos de un sistema de derechos de propiedad.

Supongamos que nuestro agricultor estima que puede criar un cerdo con un valor de mercado de 1 000 dólares a un costo de sólo 500 dólares por concepto de mano de obra y materiales, para una ganancia neta de 500 dólares, y que ningún otro uso de la tierra ocupada por el cerdo le daría un valor neto mayor: en el siguiente uso mejor, su ingreso derivado de la tierra sería sólo de 200 dólares. El agricultor deseará criar el cerdo. Pero supongamos ahora que su derecho de propiedad está restringido en dos sentidos: no tiene derecho a impedir que la máquina de un ferrocarril adyacente emita accidentalmente chispas que podrían incendiar la pocilga del cerdo, matando a éste prematuramente; y un tribunal podría decidir que la crianza de un cerdo en esta tierra es un perjuicio (véase después en III.7), en cuyo caso tendrá que vender al cerdo en términos desventajosos antes de tiempo. En vista de estas contingencias, el agricultor debe revaluar el rendimiento de su tierra. Debe descontar los 1 000 dólares estimados originalmente para reflejar la probabilidad de que el rendimiento sea mucho menor, quizá cero. Supongamos que, después de este descuento, el ingreso esperado de la crianza del cerdo (el valor de mercado multiplicado por la probabilidad de que llegue al mercado) es de sólo 600 dólares. El agricultor no criará al cerdo. La ganancia esperada es ahora de sólo 100 dólares, ya que los costos del agricultor se mantienen a 500 dólares. Dijimos que el mejor uso siguiente de la tierra daría una ganancia de 200 dólares. El agricultor le dará ese uso, aunque sea menos valioso (200 dólares contra 500 dólares), de modo que bajará el valor de la tierra.

Pero el análisis está incompleto. La eliminación del cerdo podría aumentar el valor de la tierra residencial circundante en mayor medida que la disminución del valor de la parcela del agricultor; o el costo de la prevención de la emisión de las chispas de la máquina podría superar la disminución del valor de la tierra del agricultor, cuando éste cambiara, digamos, la crianza de cerdos por el cultivo de rábanos a prueba de incendios, por ejemplo. Pero el lector sagaz podría objetar que si el aumento del valor para otros, por un uso diferente de la tierra del agricultor, supera a la disminución del valor para él, debieran estar facultados para comprar su derecho. El ferrocarril puede comprar un derecho de servidumbre o permiso de ingreso de emisiones de chispas; los propietarios circundantes pueden comprarle al agricultor un convenio para que no críe cerdos; no hay necesidad de limitar el derecho de propiedad del agricultor. Sin embargo, como veremos en III.7, los costos de la realización de una transferencia de derechos —los costos de transacción— son a menudo prohibitivos; y cuando esto ocurre, el hecho de otorgar a alguien el derecho exclusivo a un recurso puede reducir la eficiencia, en lugar de aumentarla.

Es importante tener en cuenta que la propiedad no implica alienabilidad (transferibilidad). Las personas son dueñas de sus cuerpos, pero no se les permite venderse como esclavos, vender sus órganos vitales o vender su sangre (en Inglaterra) a los bancos de sangre.[4] En capítulos posteriores analizamos algunas de estas restricciones (por ejemplo, sobre la venta de los derechos parentales). Desde un punto de vista económico, los derechos de propiedad deben ser al menos presuntivamente alienables, por lo que el mercado puede asignarlos a las personas que pueden obtener el mayor valor de los mismos.

III.2. Problemas de la creación y la aplicación de los derechos de propiedad

Los derechos de propiedad no sólo son menos exclusivos, sino también menos universales de lo que serían si su cumplimiento forzoso no fuese costoso. Imaginemos una sociedad primitiva en la que el uso principal de la tierra fuera el pastoreo. La población de la sociedad es pequeña en relación con la cantidad de tierra, y sus hatos son pequeños también. No existe ninguna tecnología para aumentar el valor de la tierra mediante fertilizantes, obras de riego u otras técnicas. El costo de la madera u otros materiales para la construcción de cercas es muy elevado, y, siendo la sociedad analfabeta, ni pensar en un sistema de registro público de la propiedad de la tierra. En esas circunstancias, los costos del cumplimiento forzoso de los derechos de propiedad bien podrían superar a los beneficios. Los costos serían los costos del cercamiento para mantener fuera a los animales de pastoreo de otras personas y serían sustanciales. Los beneficios podrían bajar a cero. Dado que no hay ningún problema de hacinamiento, los derechos de propiedad no conferirían beneficios estáticos, y, dado que no hay manera de mejorar la tierra, no habría tampoco beneficios dinámicos. No es sorprendente que el patrón en el que surgen y se desarrollan los derechos de propiedad se relacione con los aumentos de la razón de los beneficios de los derechos de propiedad a sus costos.[5]

[4] Para contrastar los puntos de vista sobre la conveniencia de limitar la alienabilidad, compárese a Susan Rose-Ackerman, "Inalienability and the Theory of Property Rights", 85 *Colum. L. Rev.*, 931 (1985), con Neil Duxbury, "Law, Markets and Valuation", 61 *Brooklyn L. Rev.*, 657 (1995).

[5] Hay una extensa literatura económica sobre el desarrollo histórico de los sistemas de derechos de propiedad: en el mundo prehistórico, primitivo y antiguo (véanse, por ejemplo, el artículo de Smith en la nota 2; Martin J. Bailey, "Approximate Optimality of Aboriginal Property Rights", 35 *J. Law & Econ.,* 183 (1992); D. Bruce Johnsen, "The Formation and Protection of Property Rights among the Southern Kwakiutl Indians", 15 *J. Legal Stud.,* 41 (1986)); en la Edad Media (véase, por ejemplo, Carl J. Dahlman, *The Open Field System and Beyond: A Property Rights Analysis of an Economic Institution* (1980)); y en el oeste norteamericano del siglo XIX (véanse, por ejemplo, Mark T. Kanazawa, "Efficiency in Western Water Law: The Development of the California Doctrine, 1850-1911", 27 *J. Legal Stud.,* 159 (1998); Terry L. Anderson y Peter J. Hill, "The Race for Property Rights", 33 *J. Law & Econ.,* 177 (1990)). Véase una síntesis en Robert C. Ellickson, "Property in Land", 102 *Yale L. J.,* 1315 (1993).

La distinción del derecho común entre los animales domésticos y los salvajes o silvestres ilustra la observación general. Los animales domésticos son propiedad como cualquier otra propiedad personal; los animales salvajes o silvestres no son propiedad hasta que sean muertos o colocados bajo una restricción efectiva, como en un zoológico. Por lo tanto, si su vaca se pasea fuera de su tierra, sigue siendo su vaca; pero si una tuza cuya madriguera se encuentra en su tierra vaga fuera de ella, no es de su propiedad, y quienquiera que la desee puede capturarla o matarla, a menos que esté domesticada —a menos, esto es, que tenga un *animus revertendi* (el hábito de volver a su tierra)—. Sería difícil hacer valer un derecho de propiedad sobre un animal salvaje; sin embargo, la excepción forjada por la doctrina del *animus revertendi* es eficiente porque fomenta la domesticación de animales salvajes o silvestres.

El caso estadunidense líder en la regla de la captura —la regla de que la propiedad de un animal salvaje se obtiene al tomar posesión de él— es el antiguo y curioso caso de *Pierson vs. Post*.[6] Post organizó una cacería de zorros. Mientras que él y los otros miembros de la caza perseguían a un zorro en una tierra sin dueño, Pierson, quien no era parte del grupo de caza, apareció y mató al zorro. La demanda era sobre quién era el dueño del zorro muerto, y el tribunal falló a favor de Pierson. Ése fue un error. Pierson fue un oportunista, y permitirle atrapar al zorro redujo la utilidad de la cacería para los cazadores.

Supongamos que, contrario al ejemplo anterior, los animales salvajes o silvestres en cuestión son valiosos. (Los zorros no se consideraban valiosos; en la caza del zorro, éste era destrozado por los perros del cazador.) Salvo que existan derechos de propiedad sobre animales valiosos por sus pieles, como la marta y el castor, los cazadores los atraparán hasta extinguirlos, aunque el valor presente del recurso disminuya al actuar así. El cazador que perdona a una castora para que pueda reproducirse sabe que los castores nacidos de ella serán atrapados casi seguramente por otro cazador (mientras haya muchos cazadores), y él no sacrificará un beneficio presente para conferir un beneficio futuro a alguien más. Los derechos de propiedad serían deseables en estas circunstancias, pero resulta difícil imaginar un plan viable para otorgar al cazador que decidiera perdonar a la castora un derecho de propiedad sobre sus crías no nacidas, a menos que comprara el hábitat del animal.[7]

[6] 3 Caines 175, 2 A.D. 264 (N.Y. Sup. Ct. 1805). Véase Dhammika Dharmapala y Rohan Pitchford, "An Economic Analysis of 'Riding to Hounds': *Pierson vs. Post* Revisited", 18 *J. Law, Econ. & Organization*, 39 (2002).

[7] Por lo que toca a la economía del comercio de pieles, véase Harold Demsetz, "Toward a Theory of Property Rights", 57 *Am. Econ. Rev. Papers & Proceedings,* 347, 351-353 (mayo de 1967), uno de los artículos señeros del "nuevo" campo de derecho y economía. Véase también Dean Lueck, "The Economic Nature of Wildlife Law", 18 *J. Legal Stud.,* 291 (1989).

De modo similar, si las zonas de desove de algunas especies de peces comestibles no pertenecen a nadie, ni están sujetas a la regulación que imita a la propiedad mediante la limitación del acceso, los pescadores podrían pescar en exceso hasta extinguir la especie. Un pescador que limita su pesca sólo estará dejando más peces para sus competidores. Pero este problema de la acción colectiva puede resolverse, y hasta cierto punto lo ha sido, mediante un acuerdo entre las naciones respecto a que cada una tendrá el control exclusivo de los derechos de pesca dentro de una distancia específica de sus costas. Dado que hay muchos menos países pesqueros que pescadores particulares, el acuerdo entre las naciones es más viable que permitir que sea el mercado el que controle la pesca excesiva.

Sin embargo, muchos de los métodos que se utilizan para evitar la sobrepesca son ineficientes comparados tanto con las tarifas para los usuarios como con las regulaciones de "límites máximos y de comercio" diseñadas de acuerdo con el sistema exitoso de permisos negociables para emitir dióxido de azufre (véase después en XIV.5). Un ejemplo es acortar la temporada de pesca, lo que podría tener poco efecto además de incentivar a los pescadores a comprar barcos más grandes para poder aumentar su captura durante la temporada acortada.

Incluso con la regulación, una especie de gran valor comercial puede extinguirse si la tasa de reproducción de la especie es menor a la tasa de descuento del mercado. Supongamos que la tasa de descuento es de 10%, pero que la tasa de reproducción es de sólo 2%. Frenar la recolección de la especie a fin de poder tener 2% más para la recolección del siguiente año no tendría ningún sentido en términos del comercio, a menos que se esperara que un aumento en la escasez eleve el precio futuro de la especie por lo menos en 8%. De otro modo, el aplazamiento de la recolección en tal caso sería similar a pedir dinero prestado con un interés de 10% para comprar un bono que sólo generaría un interés de 2 por ciento.

Los derechos de propiedad están correlacionados con la escasez. La diferencia en los sistemas de los derechos de aguas de los estados orientales y occidentales de la Unión Americana ilustra este punto. En los estados orientales, donde el agua es abundante, los derechos de aguas se hacen comunales en un grado considerable, y la regla básica es que los propietarios ribereños (es decir, los propietarios de franjas de la playa de un cuerpo de agua) tienen derecho a hacer un uso razonable del agua: un uso que no interfiera indebidamente con los de los otros propietarios ribereños. En los estados occidentales, donde el agua es escasa, pueden obtenerse derechos exclusivos mediante la apropiación (el uso).

Consideremos ahora el ejemplo de cosas —a menudo muy valiosas, como el tesoro en un barco hundido— que fueron propiedad algún día pero han sido abandonadas. Aquí la regla general es que quien las encuentre será su propietario. En cierto sentido, ésta es la misma regla que para los animales

salvajes o silvestres (incluyendo los peces que no se criaron en una granja) y para el agua en los estados occidentales. La propiedad de la cosa se obtiene reduciéndola a una posesión efectiva. Hasta que ello no ocurra, la cosa no tiene propietario (los castores no nacidos, el barco abandonado), y es esta brecha de la propiedad —el intervalo en el que nadie tiene un derecho de propiedad— lo que constituye la fuente del uso dispendioso.

Pero el problema es ligeramente diferente en el caso del animal y en el del tesoro. En el primer caso, el problema es la explotación demasiado rápida; en el segundo, la explotación demasiado costosa. Supongamos que el tesoro del barco hundido vale un millón de dólares y que costará 250 000 dólares la contratación de un equipo de buzos para que lo rescate. En virtud de que la ganancia esperada de la empresa es tan elevada, otro individuo podría decidir la contratación de su propio equipo y tratar de llegar antes que el primer equipo. Un tercero y aun un cuarto podrían intentarlo también, porque si cada uno de ellos tiene la misma probabilidad (25%) de llegar primero al tesoro, el valor esperado de la empresa para cada uno (un millón de dólares × 0.25) cubrirá todavía el costo esperado de cada uno. Con todo, si los cuatro lo intentan, el costo de la obtención del tesoro, un millón de dólares, será cuatro veces mayor de lo que sería si sólo uno lo hubiese intentado.[8] La pérdida social neta derivada de esta competencia será menor de 750 000 dólares, porque la competencia hará que el tesoro sea encontrado más pronto (lo que aumenta su valor presente) que si sólo un rescatador lo estuviese intentando. Pero el ahorro de tiempo podría ser demasiado pequeño para contrarrestar los gastos adicionales que aceleraron la búsqueda.[9]

No ocurriría tal desperdicio si el tesoro no hubiese sido abandonado; el propietario habría contratado simplemente a uno de los cuatro rescatadores por 250 000 dólares. Pero cuando hablamos de una propiedad "abandonada" en el sentido legal queremos decir que el costo de su rescate por el propietario original es prohibitivo, ya sea porque resulta imposible su encuentro a un costo razonable o porque él considera (quizás incorrectamente) que la propiedad vale menos que el costo de encontrarla o usarla, y así se ha lavado las manos. (Más adelante analizamos el abandono con mayor detalle en III.12.)

El problema de la optimización de la inversión para la obtención de los derechos de propiedad se ilustra por una comparación de tres reglas para la posesión legal de una ballena en el siglo XIX.[10] Bajo una regla, si el primer

[8] La tendencia de una ganancia esperada a traducirse en los costos mediante los esfuerzos competitivos recibe el nombre de búsqueda de renta, se la encontrará muchas veces en este libro y es el tema de una bibliografía abundante. Véanse los ensayos reunidos en *Toward a Theory of the Rent-Seeking Society* (James M. Buchanan, Robert D. Tollison y Gordon Tullock [comps.], 1980). La revista *Public Choice* publica muchos artículos sobre la búsqueda de renta.

[9] Hay una analogía con la carrera de patentes, que se discute más adelante en XI.1.

[10] Véase Robert C. Ellickson, *Order without Law: How Neighbors Settle Disputes*, 196-206 (1991).

ballenero en acertar a la ballena con su arpón no logra detenerla, no tendrá derecho a la ballena si ésta es finalmente matada por otro ballenero; ese ballenero será el propietario. Bajo una segunda regla, el primer ballenero tenía derecho a la mitad de la ballena, y bajo la tercera regla tenía derecho a toda la ballena siempre que la punta del arpón permaneciera en el animal, aunque la cuerda hubiese sido cortada. Si el derecho a la ballena le perteneciera al primer ballenero en clavar el arpón en la ballena, todo el océano podría llenarse de aficionados buenos para manejar los arpones, pero no para matar efectivamente a las ballenas. Éste sería un ejemplo de una carrera socialmente dispendiosa para ser el primer "descubridor" de una propiedad valiosa. Si, en cambio, la ley otorgara el derecho de propiedad sobre la ballena al ballenero que la matara, esto podría desalentar la actividad cooperativa, que es más importante para la eficiencia en la caza de las ballenas que en la mayoría de las actividades de caza, en las que prevalece la regla de que la propiedad puede obtenerse sólo por la posesión. La tercera solución de "la mitad de la ballena" puede entenderse como un método para alentar la cooperación.

Otra forma en la que el derecho común se ocupa de la obtención de los derechos de propiedad es a través de la doctrina de la adhesión. Si cae un meteorito en la tierra de alguien, se considera que éste es su dueño; asimismo, es dueño de las modificaciones que sus inquilinos hagan a la propiedad que les alquila, de los terrenos que se incorporen a su tierra por los cambios en la línea de la costa y de las crías de los animales que posee. La asignación de una propiedad nueva de esta manera minimiza los costos de transacción y la búsqueda de rentas.

Otra regla del derecho común que disminuye la búsqueda de renta dispendiosa en la indagación acerca de propiedades abandonadas es la regla de que el tesoro oculto y abandonado (dinero y lingotes), si se encuentra, pasa a manos del gobierno antes que convertirse en propiedad de quien lo encuentre. Esta regla reduce la inversión en la búsqueda del tesoro a cualquier nivel que el gobierno considere conveniente; el gobierno determina ese nivel fijando el monto de la recompensa que se entregará a quien lo encuentre. En el caso del dinero (por oposición al tesoro que tiene un valor histórico o estético), el nivel óptimo es muy bajo, quizá nulo. El hallazgo de dinero no aumenta la riqueza de la sociedad; solamente permite que quien lo encuentre tenga más de los bienes de la sociedad que otros. (También contribuye a la inflación al aumentar la oferta monetaria en relación con la producción total.) Por lo tanto, la recompensa óptima puede ser muy baja, quizá nula. La tendencia del derecho común es a la expansión del principio de la propiedad estatal (o principio de reversión al Estado) del tesoro oculto a otras áreas del hallazgo de propiedades, con lo cual se otorga al descubridor una recompensa antes que la propiedad misma.

Que la *negación* de un derecho de propiedad puede ser tanto un instrumento para economizar como la creación de un derecho se ilustra a detalle por

medio de la doctrina de un fideicomiso público,[11] según la cual las vías acuáticas navegables, las tierras bajas y ciertas playas quedan reservadas para el acceso público: nadie puede establecer un derecho de propiedad en ellas. Si un recurso es valioso pero no escaso (¿una paradoja?[12]), la creación de derechos de propiedad no desempeña ninguna función de economización: sólo incita a la búsqueda de renta y se traduce en la disipación del recurso. Sin embargo, otra razón de la aparición de la doctrina del fideicomiso público podría ser el deseo de limitar la búsqueda de renta prohibiendo al gobierno la cesión de propiedades públicas valiosas. En el siguiente capítulo volveremos a ocuparnos de este asunto.

III.3. Derechos de propiedad en el derecho y en la economía: frecuencias de transmisión y subastas de ondas de radio

Hasta aquí nos hemos mantenido muy cerca de la idea de un derecho de propiedad que tiene el abogado, pero a menudo divergen las concepciones legales y económicas de los derechos de propiedad. Veamos un ejemplo tomado de la transmisión.[13]

En los primeros días de la radio, antes de la regulación federal completa, había cierto apoyo judicial para la proposición de que el derecho a la transmisión en una frecuencia particular en un área particular, sin interferencia de otros usuarios, era un derecho de propiedad que podía protegerse mediante un juicio. Con la creación de la Comisión Federal de Radio (FRC), antecesora de la Comisión Federal de Comunicaciones (FCC) en el año de 1928, el Congreso adoptó una postura bastante diferente. Las concesiones que autorizaban la utilización de frecuencias particulares en áreas particulares se otorgarían, a un costo nominal, por periodos de tres años que podían renovarse, a los solicitantes que persuadieran a la Comisión de que si promoverían el in-

[11] Véanse Dean Lueck, "The Rule of First Possession and the Design of the Law", 38 *J. Law & Econ*., 393, 423-424 (1995); Lloyd R. Cohen, "The Public Trust Doctrine: An Economic Perspective", 29 *Cal. W. L. Rev.,* 239 (1993).

[12] No. Un recurso abundante podría tener un gran valor en cuanto a su uso, pero no en el *intercambio* (Adam Smith fue el primero en señalar esta diferencia); por ejemplo, el aire. Su abundancia hace que su costo de oportunidad sea de cero y, por lo tanto, en un mercado competitivo (es decir, a menos que se monopolice el suministro de aire y se genere, por consiguiente, una escasez artificial) su precio también es cero, a pesar de que es esencial para la vida.

[13] Véanse Ronald H. Coase, "The Federal Communications Commission", 2 *J. Law & Econ*., 1 (1959); Jora R. Minasian, "Property Rights in Radiation: An Alternative Approach to Radio Frequency Allocation", 18 *J. Law & Econ.,* 221 (1975); Thomas W. Hazlett, "The Rationality of U.S. Regulation of the Broadcast Spectrum", 33 *J. Law & Econ.,* 133 (1990); Dale Hatfield y Phil Weiser, "Spectrum Policy Reform and the Next Frontier of Property Rights", 15 *Geo. Mason L. Rev.*, 549 (2008).

terés público. El Congreso estableció expresamente que los concesionarios no tendrían ningún derecho de propiedad en el uso de las frecuencias que se les asignaran. El propósito principal de esta disposición era el de impedir toda reclamación de compensación por un usuario cuya concesión no fuera renovada al término del periodo de tres años.

Algunas de las objeciones que se formularon contra el reconocimiento de derechos de propiedad en frecuencias suenan feo a los oídos del economista. Por ejemplo, se dijo que, si los derechos de transmisión pudieran comprarse y venderse como otras propiedades, los medios de la transmisión caerían bajo el control de los ricos. Esto confunde la voluntad de pagar con la capacidad de pagar. La posesión de dinero no dicta los objetos que se comprarán. Los pobres les quitan a los ricos con frecuencia algunos bienes porque están dispuestos a pagar más en total.

En la administración efectiva del programa de regulación federal de la transmisión, la disposición a pagar ha desempeñado un papel decisivo y ha surgido un sistema de derechos de propiedad *de facto*. Las concesiones de radio y televisión deseables se han otorgado en procedimientos comparativos en los que la disposición a pagar —no por la licencia como tal, sino por la representación legal y la influencia política que pueden determinar el resultado— ha decidido en muchos casos quién controlará el recurso en disputa. Sin embargo, este método de asignación inicial de los derechos de transmisión es ineficiente. La competencia por la obtención de una concesión podría disipar el valor esperado de ella en gastos legales, de cabildeo y otros relacionados, tanto como en la competencia por la búsqueda de tesoros hundidos. La participación en una subasta de frecuencias radiofónicas no requeriría costosos servicios jurídicos y de cabildeo, por lo menos si puede impedirse a bajo costo la manipulación de la subasta.

Aun así, la no asignación del derecho al solicitante que lo valúe más es sólo una ineficiencia transitoria. Una vez que se han obtenido los derechos de transmisión en un procedimiento de concesión, podrán venderse como algo incidental a la venta de los activos físicos de la estación de radio o televisión. Cuando se vende por 200 millones de dólares una estación de televisión que tiene un transmisor y otras propiedades físicas que valen sólo unos cuantos cientos de miles de dólares, podemos estar seguros de que la mayor parte del precio de compra es el pago de los derechos de uso de la frecuencia. Así pues, los derechos de transmisión acaban de ordinario en las manos de quienes estén dispuestos a pagar más dinero por ellos, aunque la "subasta" inicial no haya asignado los derechos eficientemente.

La disposición de los radiodifusores a pagar decenas de millones de dólares por un derecho que termina al cabo de tres años podría parecer algo peculiar. Pero las concesiones radiofónicas se han cancelado sólo por graves errores de conducta, de modo muy similar al que nos lleva a perder nuestra tierra por no pagar los impuestos prediales.

En 1994 quedó demostrada la viabilidad de los derechos de propiedad explícitos en el espectro electromagnético cuando el gobierno subastó licencias para el uso del espectro de la telefonía móvil y otros servicios de comunicaciones personales por primera vez. Se subastaron miles de concesiones y el gobierno de los Estados Unidos se llevó en total casi ocho mil millones de dólares. La experiencia con la subasta de ondas de radio revela que el uso de los derechos de propiedad para racionar el acceso al espectro electromagnético no es sólo un sueño guajiro del economista.[14]

El concepto de un derecho de propiedad *de facto* introducido en esta sección tiene una aplicabilidad amplia. En efecto, algunos economistas emplean el término *derecho de propiedad* para describir virtualmente todo instrumento —público o privado, de derecho común o regulador, contractual o gubernamental, formal o informal— por el que se reduzcan las divergencias existentes entre los costos o beneficios privados y sociales.[15] En un libro de derecho este empleo puede resultar confuso, de modo que en general confinaremos el uso de este término a los derechos de propiedad formales.

III.4. Derechos al uso futuro

La compra de tierras baldías con la intención de conservarlas para su desarrollo futuro es un tipo de transacción común, pero la revelación de una intención de diferir indefinidamente la iniciación de operaciones, en una solicitud de concesión de transmisión, garantizaría la negativa. Lo mismo se aplica a los derechos de aguas bajo el sistema de apropiación que prevalece en los estados occidentales de la Unión Americana. Se adquieren derechos de propiedad de aguas (como de animales salvajes o silvestres) mediante la posesión, es decir, a través de la desviación y el uso efectivos de una corriente, y el derecho sólo abarca la cantidad de agua efectivamente usada; no puede obtenerse un derecho para ejercerlo más adelante. No obstante, las limitaciones de la transmisión y del agua se eluden hasta cierto punto, en el caso de la transmisión difiriendo la construcción efectiva luego de que la concesión ha sido obtenida y en el caso del agua obteniendo un permiso preliminar que establece el derecho previo del solicitante, aunque se pospongan la construcción de las obras de desviación y el uso del agua desviada.

Es posible que la hostilidad existente hacia el reconocimiento de los derechos de uso futuro se relacione con la hostilidad milenaria hacia la especulación: típicamente, la compra de un bien no con la intención de usarlo, sino

[14] Véase el Simposio "The Law and Economics of Property Rights to Radio Spectrum", 41 *J. Law & Econ.*, 521 (1998).

[15] Véanse algunos ejemplos en Armen A. Alchian y Harold Demsetz, "The Property Right Paradigm", 33 *J. Econ. Hist*., 16 (1973), y los ensayos de *The Economics of Property Rights* (Eirik G. Furubotn y Svetozar Pejovich [comps.], 1974).

para conservarlo en suspenso con la esperanza de que aumente su valor.[16] La especulación desempeña una función económica valiosa al ayudar a lograr que los precios reflejen correctamente las condiciones de la oferta y la demanda (véase también IV.11). Tiene a la vez una función estática —la de alinear los precios con la demanda y la oferta corrientes, es decir, evitar escaseces o congestionamientos presentes[17]— y una función dinámica. Un mercado de futuros (por ejemplo, un mercado en el que se puede comprar trigo a un precio fijo para su entrega no ahora sino dentro de un año) puede regular el consumo a lo largo del tiempo, cancelando los efectos de los congestionamientos y las escaseces. Por ejemplo, si se espera una escasez, los especuladores aumentarán sus compras de futuros porque esperan que el precio de mercado sea mayor el año próximo; por lo tanto, el precio de los futuros aumentará, induciendo a los vendedores a guardar una porción de la oferta corriente, o a cobrar un precio mayor a los compradores corrientes —pero se tendrá el mismo efecto si se reduce la demanda y, por lo tanto, la oferta actual—, a fin de aprovechar el precio mayor esperado en el futuro. La expansión resultante de la oferta futura moderará a su vez ese precio; en esta forma, la especulación puede reducir las fluctuaciones de los precios.

Las compras para uso futuro no son necesariamente especulativas; pueden ser todo lo contrario: un instrumento de protección. Un agricultor podría saber que necesitará más agua para riego dentro de pocos años y, en lugar de correr el riesgo de los cambios de precios del agua, puede decidir firmar un contrato ahora, a un precio fijo, para la entrega futura de una cantidad especificada de agua. (Por lo tanto, el vendedor estaría especulando con los cambios futuros del precio del agua: ¡la especulación facilita la protección!) Si se prohíben tales ventas, el agricultor podría optar por usar ahora más agua de la que realmente necesita, sólo para asegurarse de tener un derecho al agua en el futuro, cuando la necesitará. El efecto principal de la prohibición de compras de agua o de frecuencias de transmisión o de bancos de ostras para uso futuro es un estímulo a los usos no económicos, no destinados a satisfacer la demanda sino a establecer un derecho.

[16] O la venta de un bien con la esperanza de que su valor baje, como ocurre con la venta de acciones en corto o de otros instrumentos financieros. En este párrafo se da un ejemplo y en el capítulo XVI se presenta un análisis más detallado.

[17] Para ilustrarlo, supongamos que una ferretería almacena lámparas de mano en previsión de un huracán. El huracán llega y, aunque la tienda ha aumentado el precio de las lámparas de mano muy por encima de su nivel normal, pronto se venden todas las lámparas menos una. Aparece entonces otro cliente que desea la lámpara para su hija, a quien le asustan las tormentas. El dueño de la tienda fija un precio muy elevado y el cliente desiste de la compra. Pocos minutos más tarde, llega un rescatista profesional a comprar la última lámpara a fin de buscar personas atrapadas por los destrozos causados por el huracán. Gracias al alto precio "especulativo" cobrado por el tendero, una lámpara de mano está disponible para el uso de mayor valor del rescatador, lo que no habría ocurrido si el tendero se hubiese negado a especular con la tendencia del precio.

El problema del uso prematuro es analíticamente similar al de la inversión excesiva en el hallazgo de tesoros escondidos. A fin de adquirir un derecho valioso, la gente puede invertir recursos más allá del punto en que tales recursos confieren un beneficio social neto. Construir una casa y cultivar un terreno en tierras públicas es otro ejemplo.[18] Si se otorga una tierra sin un cobro pecuniario, pero el regalo se condiciona (como en efecto ocurrió con la Ley del Homestead de 1862) a la ocupación efectiva de la casa y el cultivo de la tierra, el residente la cultivará hasta el punto en que el último dólar del esfuerzo rinda un dólar de beneficio en la obtención de este derecho, aunque se coseche menos de un dólar de producción agrícola. Pero si el gobierno *desea* que la tierra sea ocupada a la mayor brevedad posible por razones políticas o militares, una ley del *homestead* podría ser un método eficiente para la realización de este propósito.[19]

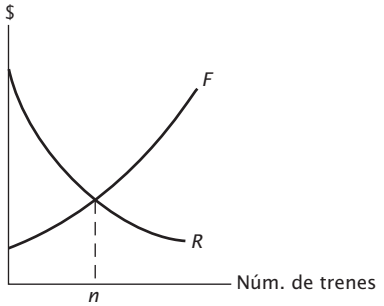
III.5. Usos incompatibles; una revisión del Teorema de Coase

Los derechos de propiedad verdaderamente exclusivos (absolutos, sin calificaciones) serían una contradicción en los términos. Si un ferrocarril ha de disfrutar el uso exclusivo de su derecho de vía, debe permitírsele que su máquina emita chispas sin limitación legal; de otro modo, el valor de su propiedad bajará. Pero el valor de la tierra de cultivo adyacente disminuirá a causa del riesgo de incendio por las chispas. ¿Debería ser juzgada la emisión de chispas como un incidente del derecho de propiedad del ferrocarril (es decir, una parte de su conjunto de derechos), o como una invasión del derecho (o conjunto de derechos) de propiedad del agricultor?

Recordemos que el Teorema de Coase, mencionado brevemente en el capítulo I, instruye que si los costos de transacción son nulos (de modo que la transferencia de un derecho de propiedad no genera ningún costo), la asignación inicial de recursos se revertirá si la revocación aumenta la eficiencia de la asignación. La operación del Teorema de Coase se representa en la gráfica III.1. La curva *R* muestra el ingreso marginal del ferrocarril como una función del número de trenes que corren cada día. La curva es descendente porque se supone que cada tren adicional contribuye con menos ingreso neto que el tren anterior. La curva *F* muestra el costo marginal del daño sufrido por el cultivo del agricultor, también como una función del número de trenes. Esta curva asciende a medida que aumenta el número de trenes, bajo el supuesto

[18] Véase, por ejemplo, Terry L. Anderson y Peter J. Hill, "The Race for Property Rights", 33 *J. Law & Econ*., 177 (1990). Véase una exposición general en David D. Haddock, "First Possession *versus* Optimal Timing: Limiting the Dissipation of Economic Value", 64 *Wash. U. L. Q.,* 775 (1986).

[19] Douglas W. Allen, "Homesteading and Property Rights; or, How the West Was Really Won", 34 *J. Law & Econ.,* 1 (1991).



GRÁFICA III.1

de que el agricultor puede ajustarse a algún daño derivado de las chispas, pero cada tren adicional es más destructivo que el precedente. Se supone que el cambio del número de trenes es el único procedimiento para modificar el monto del daño para el cultivo del agricultor. Independientemente de que el ferrocarril sea o no responsable de ese daño, el número de los trenes operados cada día será *n* si los costos de transacción son nulos. A la izquierda de *n*, el ferrocarril puede, corriendo trenes adicionales, añadir más a sus ingresos que a los costos del agricultor, de modo que por supuesto el ferrocarril aumentará el número de trenes. A la derecha de *n*, donde la disminución del número de trenes añadiría más a los ingresos netos del agricultor que lo que restaría a los ingresos del ferrocarril, el agricultor pagará al ferrocarril para que disminuya el número de trenes hasta *n*. El número será el mismo si el agricultor tiene el derecho a verse libre de los daños al cultivo, en lugar de que el ferrocarril tenga el derecho de emitir chispas. A la derecha de *n*, el agricultor demandará al ferrocarril para que disminuya el número de trenes; a la izquierda de *n*, el ferrocarril pagará al agricultor para que renuncie a una porción de su derecho a librarse del daño.

No obstante, Coase también formula tres observaciones importantes acerca del caso en que los costos de transacción son tan elevados que no resulta viable una transferencia voluntaria de propiedad:

1. Es posible que, si se asigna la responsabilidad a la parte que en algún sentido crudo "causa" el daño, es decir, la parte activa (el ferrocarril, en nuestro ejemplo), no se produzca la solución eficiente para el conflicto. La parte pasiva puede ser (en una frase acuñada por Guido Calabresi) la "que evita el daño al menor costo". Supongamos que el agricultor tiene el derecho de propiedad y que éste vale menos que el costo para el ferrocarril de no ser capaz de correr sus trenes; sin embargo, debido a los altos costos de transacción, el ferrocarril no puede comprar la propiedad del agricultor (quizás hay demasiados agricultores que poseen tierras contiguas al ferrocarril, y si sólo uno se negara a vender

podría obtener un mandato judicial para que el ferrocarril no corriera sus trenes en las vías contiguas a su propiedad). En ese caso sería eficiente asignar el derecho de propiedad al ferrocarril, en primer lugar, antes que al agricultor.

2. El derecho común del perjuicio (véase después en III.7) puede entenderse como un esfuerzo por mejorar el destino de los recursos mediante la asignación del derecho de propiedad a la parte de un conflicto sobre el uso de la tierra para quien tal derecho sería más valioso.
3. Al decidir si es apropiada la intervención gubernamental en el sistema económico, no basta demostrar que el mercado operaría imperfectamente sin ella; el gobierno también opera imperfectamente. Lo que se requiere es una comparación entre el funcionamiento efectivo del mercado y el del gobierno en el contexto particular. Coase creía que el hecho de que las partes pudieran transar acerca de una asignación del derecho de propiedad o una regla de responsabilidad ilustraba la flexibilidad de los mercados, mientras que la tendencia del gobierno a concentrar sus esfuerzos reguladores exclusivamente en la parte activa de una interacción nociva (por ejemplo, requiriendo que los ferrocarriles equiparan sus locomotoras con guardachispas) ilustraba la rigidez de buena parte de la regulación gubernamental.

Los costos de transacción nunca son nulos y tienden a aumentar de manera abrupta a medida que aumenta el número de partes en una transacción; la fórmula para el número de conexiones necesarias para unir a todos los elementos de un conjunto de n miembros es sugestiva: $n(n-1)/2$. Sin embargo, se puede dar contenido empírico al Teorema de Coase al replantearlo como la hipótesis de que la asignación inicial de derechos de propiedad no afectará el uso final de una propiedad si se permiten las transacciones y no son demasiado costosas en relación con los beneficios de la reasignación. Se han hecho esfuerzos para comprobar la hipótesis, con resultados variados.[20]

Incluso cuando los costos de transacción son lo suficientemente bajos para que el Teorema de Coase se mantenga como una aproximación, la asignación inicial de derechos puede afectar a la eficiencia. Dado que las transacciones tienen siempre un costo en el mundo real, se promueve la eficiencia asignando el derecho a la parte que lo compraría —el ferrocarril en nuestra primera situación hipotética y el agricultor en la segunda— si estuviera asignado inicialmente a la otra parte. Y, evidentemente, el costo de la transacción es tan elevado en relación con su valor que resulta antieconómica; la asignación inicial de los derechos es definitiva y, por lo tanto, crucial.

[20] Este y otros aspectos del Teorema de Coase son el objeto de una bibliografía extensa, gran parte de la cual se recopila en *The Coase Theorem* (2 volúmenes; Francesco Parisi y Richard A. Posner [comps.], 2013).

El derecho común cuenta con reglas diseñadas para evitar una asignación económica aleatoria de los derechos de propiedad. Por ejemplo, bajo el derecho común inglés, se consideraba que un terrateniente que construyera una estructura que obstruyese la ventana de su vecino de tal modo que el vecino necesitara luz artificial para leer en la mitad del cuarto más cercana a la ventana estaría infringiendo los derechos de propiedad del vecino, siempre que éste hubiese tenido un acceso libre a la luz durante 20 años (esto alentó una preferencia por algunos sitios para la construcción y por el número y tamaño de las ventanas que minimizaría los casos de litigios por obstrucción). Consideremos las consecuencias si el derecho de propiedad se hubiese otorgado, por el contrario, a la parte que construyó después (B); supongamos entonces que el costo para la persona cuyas ventanas fueron obstruidas (A) excedería al costo de B para hacer retroceder ligeramente su pared (todo eso sería necesario, dado lo limitado del derecho); B compraría el derecho. Así, la asignación del derecho a esta persona en primera instancia evita la transacción y sus costos correspondientes. Pero los tribunales no extendieron la regla para proteger las vistas distantes. Si A tuviera una casa en una colina con una vista hermosa, y B construyera una casa que arruinara esa vista, A no podría quejarse de una invasión de sus derechos de propiedad, aunque el valor de su propiedad hubiese disminuido. La presunción de los valores relativos se revierte en este caso. Dado que la casa con una vista tiene un área de terreno mayor, es probable que los valores que se crearían con el desarrollo de tal área fueran mayores que la pérdida de valor para el terrateniente cuya vista se ve perjudicada por el desarrollo.

Esta regla del derecho común inglés ("servidumbre de vista") fue rechazada en Estados Unidos, donde la tierra era tan abundante en relación con la población que el riesgo de que la construcción bloqueara el acceso de los vecinos a la luz era mínimo.

Otra regla del derecho común (convencionalmente una regla del derecho aplicable a los cuasidelitos, pero, en virtud de que define los derechos del dueño de la propiedad, puede verse también como una regla del derecho de propiedad) era que un ferrocarril no tenía ninguna obligación de cuidar a las personas que usaran las vías como caminos, excepto en los cruces. El costo del uso de caminos alternativos sería generalmente pequeño en comparación con el costo para el ferrocarril de volver seguras las vías para quienes quisieran cruzarlas. Sin embargo, el derecho del ferrocarril era calificado: el ferrocarril estaba obligado a tener cuidado con el ganado.[21] A los agricultores les resultaría muy costosa la construcción de cercas que con seguridad impidieran las fugas del ganado, de modo que, si fuesen viables las transacciones entre los agricultores y los ferrocarriles, los agricultores pagarían con frecuencia a los ferrocarriles para que tuvieran cuidado con los animales en la vía.

[21] Véase 3 Byron K. Elliott y William F. Elliott, *A Treatise on the Law of Railroads,* 1205 (1897).

Como en el caso del ganado, la carga de la prevención de accidentes con niños allanadores, en ausencia de una obligación del terrateniente, recaería en el custodio adulto y sería costosa de sobrellevar: los niños son difíciles de encerrar. Pero incluso los niños pequeños tienen más conciencia que las vacas o las ovejas. La doctrina del perjuicio atractivo provee una solución ingeniosa para los intereses en conflicto de padres y terratenientes. El terrateniente debe cercar o proteger de algún otro modo, contra los niños allanadores, las condiciones artificiales de la tierra (clásicamente, las plataformas giratorias de los ferrocarriles) que los niños pequeños toman erradamente por juguetes inocuos. Sería impracticable que el niño o sus padres se protegieran completamente contra este tipo de riesgo, mientras que el terrateniente puede hacerlo a un costo relativamente pequeño: en todo caso, menor que el costo del cercamiento de todo el derecho de vía del ferrocarril. Ésta es, por cierto, otra situación en la que la asignación inicial de los derechos es también la asignación final; sería impracticable que los terratenientes negociaran por adelantado con todos los padres cuyos niños podrían meterse a su propiedad.

La teoría económica implica que los derechos de propiedad se redefinirán de tiempo en tiempo a medida que cambian los valores relativos de los usos diferentes de la tierra. El cercamiento del ganado provee también una ilustración. Supongamos que el ganado se sale de la tierra donde pasta y se mete a la tierra de un vecino, donde causa daños a los cultivos. ¿Debiera el vecino soportar el costo, según la teoría de que debería haber cercado su terreno, o debiera soportarlo el dueño del ganado según la teoría de que el dueño debe cercar su ganado? La respuesta pareciera depender (y una comparación de las reglas a lo largo del tiempo y entre diferentes jurisdicciones del derecho común sugieren que sí depende)[22] de la razón del ganado-cultivos. Si se destina más tierra a pastos que a cultivos, será más barato que los agricultores cerquen su tierra y no que los ganaderos cerquen las suyas, y la ley echará la carga del cercado sobre los hombros de los agricultores. La carga se revertirá cuando la razón de los usos de la tierra se invierta.

¿Le preocupa que la redefinición continua de los derechos de propiedad para obtener la eficiencia bajo condiciones cambiantes pudiera crear inestabilidad y desalentar la inversión? X compra una granja mucho tiempo antes de que haya un ferrocarril en su área. El precio de compra no se descuenta para reflejar el daño que causarán en el futuro las chispas a los cultivos, porque no se prevé la construcción de una línea de ferrocarril. Pero finalmente se construye una línea lo bastante cercana a la granja de X para que las chispas causen daños. X demanda al ferrocarril, pero el tribunal falla que el nivel de emisión de chispas es razonable porque sería más costoso que el ferrocarril impidiera la pérdida de los cultivos. Estando los valores de la propiedad expuestos así a una depreciación sin compensación por causa de los cambios

[22] Véase Richard A. Posner, *Tort Law: Cases and Economic Analysis,* 492-493 (1982).

imprevistos en los usos de la tierra vecina, disminuirá el incentivo para invertir en la agricultura.[23] Pero, como vimos en nuestro ejemplo anterior de la crianza de cerdos (III.1), la disminución del nivel de la inversión en la agricultura podría ser un ajuste eficiente ante la posibilidad de que algún día el valor más alto de la tierra del agricultor pueda ser como campo de descarga para la caída de chispas del ferrocarril.

No obstante, cuando los derechos de propiedad están sujetos a verse redefinidos a medida que cambian los valores, se impone desutilidad para las personas que sienten aversión por el riesgo. Y puede dudarse de que cualquiera de los métodos de eliminación de los riesgos sea viable.[24] Sin embargo, esto puede no ser un problema grave. Si se prevé un uso dañino de la tierra vecina en el momento de la venta, el precio de la tierra disminuirá en consecuencia, y el comprador no tendrá expectativas frustradas. Si no se prevé ese uso es probable que se le encuentre más allá en un futuro, y un costo en el que se incurra en el futuro lejano tendrá escaso impacto sobre las decisiones presentes, a menos que sea astronómico (véase después VI.7). La alternativa de asignar siempre el derecho de propiedad al más antiguo de dos usos de la tierra en conflicto sería muy ineficiente; el uso más nuevo será a menudo el más valioso, aunque los costos de transacción podrían ser prohibitivos. Esto se da además de que el problema de la regla de primero en tiempo, primero en derecho puede llevar a comprometer los recursos de manera prematura o excesiva.

Varios estados de la Unión Americana han promulgado leyes de "derecho a la agricultura" que prohíben la consideración de un uso agrícola como un perjuicio.[25] El único argumento económico en favor de tales leyes —que a primera vista transgreden la eficiencia— es que en su ausencia podrían los vecinos invertir recursos para persuadir a las juntas de zonificación o los tribunales para que declaren como un perjuicio a un uso agrícola aunque el beneficio de ese uso supere a los costos de los vecinos. Por supuesto, tales gastos en la búsqueda de renta son posibles (vimos un ejemplo cuando discutimos las frecuencias radiofónicas), como lo son las decisiones erróneas no inducidas por tales gastos, pero éstas parecen razones rebuscadas para eximir a una actividad de toda responsabilidad por sus costos externos.

[23] Pero el ejemplo del texto no es realista, ya que la llegada del ferrocarril aumentaba de ordinario el valor agrícola de la tierra cercana.

[24] Es posible que el agricultor no tuviera la posibilidad de eliminar el riesgo ya sea diversificando (por ejemplo, adquiriendo grandes cantidades de acciones comunes del ferrocarril) o comprando un seguro. Sería difícil comprar un seguro contra una disminución de los valores de la tierra causada por los cambios de la definición de los derechos de propiedad, porque la prima apropiada, que depende de la probabilidad y la magnitud de la pérdida si se materializa el riesgo, sería muy difícil de computar (véase después IV.5).

[25] Véanse Keith Burgess-Jackson, "The Ethics and Economics of Right-to-Farm Statutes", 9 *Harv. J. Law & Pub. Policy,* 481 (1986); Jo Kwong y John Baden, "Comment: The Ethics and Economics of Right-to-Farm Statutes", 9 *Harv. J. Law & Pub. Policy,* 525.

III.6. El allanamiento y el dominio eminente

El derecho del terrateniente para repeler una intrusión física bajo la forma de chispas de máquina es un derecho calificado. El intruso puede derrotarlo demostrando que su uso de la tierra, que es incompatible con el uso del terrateniente afectado, es más valioso. Pero si su vecino estaciona su automóvil en su cochera tiene derecho a echarlo como un allanador, por más convincentemente que pueda demostrar ante un tribunal que el uso de su cochera para estacionar su automóvil es más valioso que su uso de la misma cochera.

La diferencia entre los casos, por lo menos en un primer examen del problema, es la diferencia que existe entre las *reclamaciones* en conflicto y los *usos* en conflicto. En general, el método adecuado (por ser más barato y más correcto) para la resolución de los conflictos de reclamaciones es el mercado, esto es, la transacción voluntaria. Si tu vecino cree que tu cochera vale más para él que para ti, podrá pagarte para que se la rentes. Pero si sólo *reclama* que puede usar tu cochera más productivamente y, con base en ello, se hace de ella y te reta a probar lo contrario, impone a los tribunales una difícil cuestión probatoria: ¿quién de ustedes estaría *realmente* dispuesto a pagar más por el uso de la cochera? En el caso de las chispas, la negociación por adelantado podría ser inviable a causa del número de propietarios potencialmente afectados, de modo que si los tribunales quieren estimular el uso más productivo de la tierra no podrán evitar la comparación de los valores de los usos rivales. Ése no es necesariamente el caso de la cochera porque las transacciones son viables; por lo que la ley prohíbe la toma, haciendo obligatoria la negociación.

Pero si el *gobierno* quiere usar mi cochera, puede confiscarla bajo el poder del dominio eminente, pagándome una "compensación justa" (= valor de mercado); no necesita negociar conmigo. Este resultado es inconsistente con la distinción que acaba de sugerirse, porque es un caso de reclamaciones rivales antes que de usos rivales. El argumento familiar de que el poder de dominio eminente es necesario para superar la terquedad de personas que se niegan a vender a un precio "razonable" (es decir, de mercado) es mala economía. Por lo general, para el dueño de una propiedad su valor es mayor al valor de mercado —es por eso que es propiedad de éste y no de otra persona— porque se ajusta más a sus gustos o necesidades como consecuencia de su ubicación o mejoras o porque los costos de reubicación serían altos; el efecto de la dotación también constituye un factor (véase I.4 antes). Los bienes raíces son un bien heterogéneo, por lo que una parcela determinada en manos de un propietario particular generalmente le dará un valor idiosincrásico que está por encima del valor de mercado. El dominio eminente interviene para confiscar ese valor mediante impuestos; si el valor de mercado de una parte de una propiedad es de $X y el valor total (incluyendo el valor

idiosincrásico) es de $1.2X, entonces si el gobierno confisca esa parte de la propiedad bajo el dominio eminente prácticamente pagará por ella gastando $X de sus propias arcas y $0.2X del bolsillo del propietario.

Un buen argumento económico en favor del dominio eminente, aunque tiene mayores aplicaciones a los ferrocarriles y otras compañías con derechos de vía que al gobierno, es que hay necesidad de prevenir el monopolio. Una vez que el ferrocarril o el gasoducto han empezado a construir su línea, el costo de su abandono por una ruta alternativa se vuelve muy elevado. Sabiendo esto, los propietarios de las tierras que se encuentran en la ruta de la línea en avance se sentirán tentados a exigir un precio muy alto: un precio mayor al costo de oportunidad de la tierra. (Éste es un problema de monopolio bilateral: véase después III.7.) Los costos de transacción serán elevados, los costos de adquisición de la tierra también, y por ambas razones la compañía con derechos de vía tendrá que aumentar el precio de sus servicios. El precio mayor inducirá a algunos consumidores a cambiarse a servicios sustitutos. Las compañías con derechos de vía tendrán entonces una producción menor, de modo que necesitarán y comprarán menos tierra de la que hubieran comprado a precios iguales a los costos de oportunidad de la tierra. Los precios más altos de la tierra darán también a las compañías un incentivo para sustituir algo de la tierra que habrían comprado por otros insumos. A resultas de todo esto, la tierra que habría sido más valiosa para una compañía con derechos de vía que para sus propietarios actuales permanecerá en sus usos actuales, menos valiosos, y esto es ineficiente. Existen métodos mediante los cuales la compañía con derechos de vía podría invalidar el poder de retenedores que tienen los terratenientes, como veremos más adelante, pero son costosos.

Si no se tuviera que pagar una compensación justa por la confiscación, el gobierno tendría un incentivo para sustituir con la tierra otros insumos que serían más baratos para la sociedad en conjunto, pero más caros para él, una sustitución que sería ineficiente. Supongamos que el gobierno puede elegir entre la construcción de un edificio alto pero estrecho en un terreno pequeño y un edificio bajo pero amplio en un terreno grande. El valor de mercado del terreno pequeño es de un millón de dólares y el del terreno grande es de tres millones de dólares. La construcción del edificio alto y estrecho costaría 10 millones de dólares, y la del edificio bajo y amplio costaría 9 millones de dólares. La alternativa más barata desde el punto de vista de la sociedad en conjunto es la construcción del edificio alto en el terreno pequeño (costo total: 11 millones de dólares [10 millones + 1 millón]), antes que la construcción del edificio bajo en el terreno grande (costo total: 12 millones de dólares [9 millones + 3 millones]). Pero si la tierra no le cuesta nada al gobierno, construirá el edificio bajo en el terreno grande, porque entonces su costo neto será un millón de dólares menor (10 millones – 9 millones). Esto supone que el gobierno toma sus decisiones de adquisición aproximadamente como lo haría un em-

presario privado, es decir, sobre la base de los costos privados antes que los sociales, a menos que se vea obligado a tomar en cuenta éstos. El supuesto es realista; el gobierno es sensible al gasto presupuestario. Los estudiosos de los asuntos militares han señalado desde hace mucho tiempo que un reclutamiento, que vuelve artificialmente baratos los insumos de trabajo para los militares, induce una sustitución excesiva del capital por el trabajo en la función de producción militar (véase después XVIII.2). Ocurriría la misma clase de distorsión en el uso de la tierra como un insumo en los servicios gubernamentales si no hubiese ningún requerimiento de compensación por la confiscación de la tierra.

Este análisis demuestra que no es fundamental la distinción que se establece entre las reclamaciones de un recurso en conflicto y los usos en conflicto o incompatibles de los recursos. Lo que es fundamental es la distinción que se establece entre los contextos de bajos costos de transacción y de altos costos de transacción. En el primer caso, la ley debiera requerir que las partes transijan en el mercado; puede hacerlo volviendo absoluto (o casi) el derecho de propiedad del dueño actual, de modo que quienquiera que piense que la propiedad vale más tenga que negociar con el dueño. Pero en los contextos de altos costos de transacción debe permitirse que los individuos recurran a los tribunales para cambiar los recursos a un uso más valioso, porque el mercado es por definición incapaz para desempeñar esta función en tales contextos. Esta distinción se refleja en la ley sólo imperfectamente. Aun cuando algunas confiscaciones gubernamentales de tierras ocurren en contextos de altos costos de transacción —la confiscación de tierras para una carretera, o para un aeropuerto o una base militar, que requiere la reunión de gran número de parcelas contiguas—, muchas otras no lo hacen (escuelas públicas, oficinas postales, edificios de oficinas gubernamentales).

La compensación por confiscaciones forzadas es fundamental, pero, como mencionamos antes, la "compensación justa" no es una compensación total. En consecuencia, el ejercicio del poder de dominio eminente es una forma de la tributación que confisca valores subjetivos (a los que referí anteriormente como idiosincrásicos) por medio de impuestos. Como veremos en el capítulo XVIII, el mejor impuesto es aquel que no cambia el comportamiento de los individuos gravados; y dado que la incidencia del poder de dominio eminente resulta difícil de pronosticar, el "impuesto" del dominio eminente podría estar muy bien desde este punto de vista. Pero es ineficiente en otro sentido. A diferencia del impuesto habitual, el cual toma un poco de muchas fuentes, el "impuesto" de dominio eminente toma mucho de pocas fuentes. Los valores subjetivos asociados a la propiedad de una casa particular podrían ser una fracción considerable de nuestra riqueza, y su pérdida podría no estar asegurada, porque no se pueden calcular; la compañía de seguros temería que el dueño de la casa se sobreasegurara, lo que crea el problema conocido como "azar o riesgo moral", el cual veremos más adelante en IV.6.

Para una persona que sienta aversión por el riesgo, como ocurre con la mayoría de nosotros, el riesgo de perder una fracción considerable de nuestra riqueza no se verá compensado por el ahorro de costos del gobierno, aunque tal ahorro se traslade al público bajo la forma de impuestos más bajos. La pérdida de utilidad para el propietario que siente aversión por el riesgo es un costo del "impuesto" del dominio eminente hipotético y, en total, podría volverlo ineficiente.[26] Raras veces habría ganancias súbitas compensatorias, ya que una persona que valúa su propiedad menos que el mercado la venderá antes que esperar a que el gobierno la declare como un bien público.

Sin embargo, la compensación por la destrucción de los valores subjetivos mediante una confiscación bien podría ser un error, aunque tales valores pudieran determinarse correctamente a bajo costo.[27] Podría inducir a una inversión excesiva en las propiedades cuyo dueño tuviera razones para pensar que probablemente serían confiscadas tarde o temprano por el gobierno. La ley trata de resolver este problema negando la compensación por cualesquiera mejoramientos de la propiedad hechos después del anuncio de la intención del gobierno de confiscarla. El problema es más grave entre más generosa sea la compensación mínima que se considere justa.

Otra exclusión de la compensación justa que podría tener sentido práctico es la negativa de la compensación del prestigio comercial cuando se confiscan instalaciones empresariales. Aquí el problema no es tanto la medición (aunque los tribunales afirman lo contrario) como la incertidumbre acerca de que el prestigio comercial esté realmente ligado a las instalaciones; si puede transferirse intacto a otras instalaciones, no habrá sido tomado junto con la tierra.

Surgen algunos problemas difíciles cuando el valor de mercado de una propiedad confiscada por el gobierno aumenta a causa del gobierno mismo. Supongamos que en tiempos de guerra el gobierno requisa gran parte de los barcos de propiedad privada del país, y la enorme disminución de la dotación de barcos en el mercado privado eleva el precio de mercado de los barcos restantes, los que no fueron requisados.[28] ¿Debiera pagar el gobierno el nuevo precio de mercado por cualesquiera requisas posteriores? Si la respuesta es afirmativa, el resultado será una caprichosa distribución de la riqueza de los contribuyentes a los propietarios de botes. Pero la respuesta negativa es también problemática: hará que el gobierno requise demasiados barcos en virtud de que no considerará las necesidades rivales de los restantes demandan-

[26] Lawrence Blume y Daniel I. Rubinfeld, "Compensation for Takings: An Economic Analysis", 72 *Cal. L. Rev*., 569 (1984).

[27] Quizá pudiera hacerse tal cosa. La antigua Atenas tenía un método sagaz de autoevaluación de la propiedad para fines tributarios. Cualquiera podía obligarlo a uno a venderle su propiedad al valor estimado por uno mismo. Una medida simple del límite inferior del valor total (objetivo más subjetivo) de la propiedad confiscada por el dominio eminente sería el rechazo reciente, por parte del autor, de una oferta hecha de buena fe a un precio superior al del mercado.

[28] Véase United States *vs.* Cors, 337 U.S., 325, 333 (1949).

tes privados de barcos. Con todo, la respuesta afirmativa es peor, porque es probable que la mayoría de las propiedades requisadas por el gobierno se hayan beneficiado del gasto gubernamental. De hecho, toda la tierra de propiedad privada se beneficia de los gastos públicos en el mantenimiento del orden público, en un sistema de registro de los títulos, etc. Sin embargo, es posible que los beneficios se hayan incluido desde largo tiempo atrás en el precio de la tierra, en cuyo caso el pago de una compensación completa (es decir, sin la sustracción de la contribución del gobierno al valor de la propiedad) no conferirá una ganancia inesperada al propietario actual.[29]

El fallo de la Corte Suprema en el caso *Kelo*[30] plantea la cuestión de si es siempre adecuado para el gobierno usar el poder de dominio eminente para confiscar la propiedad privada a fin de transferirla a otra entidad privada, antes que conservarla para algún uso gubernamental, como una oficina postal o una base militar. La Quinta Enmienda a la Constitución permite que el gobierno confisque la propiedad sólo para "uso público". En el caso *Kelo*, Pfizer, la empresa farmacéutica, había decidido construir un centro de investigación grande junto a un trecho adyacente de 90 acres (36 hectáreas) de una propiedad que da al centro de la ciudad en la línea costera de New London. La ciudad esperaba que la presencia de Pfizer atrajera a otras empresas al vecindario. Los demandantes poseían propiedades residenciales ubicadas en secciones del trecho de 90 acres que el plan de reurbanización de la ciudad reservó como espacios para oficinas y estacionamientos. Es probable que el desarrollo de esas áreas para tal uso fuera imposible si en éstas hubiera casas por un lado y por otro (los demandantes poseían 15 casas en total entre las dos áreas). La ciudad llamó a esto el problema de las "manchas de leopardo" y lo resolvió declarando el uso público de las casas. Sin embargo, no defendió su acción en función del poder de retención de los demandantes, aunque la analogía con la declaración de uso público de una empresa con derecho de vía es clara. En lugar de ello, señaló que "el área [del proyecto de reurbanización, una zona costera en el centro de New London, Connecticut] estaba en suficientes apuros económicos para justificar el programa de rejuvenecimiento económico". Pero si el "rejuvenecimiento económico" es un uso público, ¿qué habrá de evitar que la ciudad confisque los hogares de las familias de clase media baja y se los dé a los multimillonarios de manera gratuita, siempre que pudiera demostrar que los nuevos propietarios estarían dispuestos a pagar más por diversos bienes y servicios locales, y por la

[29] Supongamos que A compró la propiedad por mil dólares y la vendió a B por dos mil dólares debido a la actividad del gobierno subsecuente. Si el gobierno declara la propiedad como pública debería pagar dos mil dólares a B. El beneficio derivado de la actividad gubernamental ha sido una ganancia súbita para su antecesor, pero no para B, quien pagó en su totalidad el precio que había pagado por la propiedad.

[30] Kelo *vs*. City of New London, 549 U.S. 465 (2005). Véase Abraham Bell y Gideon Parchomovsky, "The Uselessness of Public Use", 106 *Colum. L. Rev.*, 1412 (2006).

propiedad y otros impuestos locales, para contrarrestar el gasto de compensación a los dueños de las propiedades confiscadas al valor de mercado? Al ratificar la confiscación, la Corte Suprema hizo caso omiso del hecho de que los constructores privados que desean montar una gran parcela contigua de tierra generalmente lo pueden hacer valiéndose de compradores testaferros. El hecho de que para el gobierno sea difícil intervenir con la discreción que requiere el método del testaferro para que funcione es un mal argumento para permitir que el gobierno utilice el dominio eminente a favor de los constructores privados antes que dejar que se las arreglen por sí solos.[31]

Un argumento mejor que justifica la decisión de la Corte es que mientras más se limite la urbanización privada de las tierras confiscadas, el mismo gobierno se volverá más activo en la urbanización, y eso sería ineficiente. Si la ciudad de New London hubiera construido un área de oficinas, de estacionamiento, etc., en los terrenos que se confiscaron a los propietarios privados, sería poco probable que una impugnación basada en la limitación del "uso público" tuviera éxito, a menos que la Corte restringiera el "uso público" de manera terminante para demorar situaciones y estuviera preparada para tratar de determinar, caso por caso, si existe tal situación genuina.

Analicemos ahora lo que se debe considerar como una "confiscación". Se ha argüido que, si el principio de la compensación justa se fundara en consideraciones de la eficiencia, y si el valor de mercado de mi casa bajara en 10 000 dólares a resultas de alguna regulación gubernamental, yo tendría derecho a la misma compensación que si el gobierno hubiese tomado una esquina de mi propiedad que vale 10 000 dólares.[32] Pero hay algunas diferencias económicas notables entre estos dos casos. Cuando una regulación gubernamental que afecta los valores de las propiedades es general en su aplicación, como ocurrirá de ordinario, los costos de la realización de la compensación serían muy elevados; esto sería especialmente probable si se hicieran esfuerzos, como en la lógica económica debieran hacerse, para tomar en cuenta a las personas beneficiadas por la regulación, otorgándoles una compensación negativa (es decir, gravando sus ganancias inesperadas). Imaginemos las dificultades que afrontaría el gobierno para identificar y luego transar con cada individuo cuya propiedad aumentara o disminuyera de valor por la regulación gubernamental del precio del gas natural o del petróleo para calefacción. Además, dado que una regulación afecta a más individuos que una sola confiscación, es más probable que movilice a una oposición política efectiva. Un freno político es menos probable, incluso en el caso de una serie aleatoria de confiscaciones (por oposición a una sola confiscación aislada), porque es menos probable que las víctimas constituyan un grupo homogéneo capaz de movilizarse efectivamente para la acción política.

[31] Véase Daniel B. Kelly, "The 'Public Use' Requirement in Eminent Domain Law: A Rationale Based on Secret Purchases and Private Influence", 92 *Cornell L. Rev.*, 1 (2006).

[32] Bruce A. Ackerman, *Private Property and the Constitution*, cap. 6, 1977.

Interviene una consideración adicional cuando la regulación afecta a usos de la tierra interactivos. Un ejemplo sería el de un decreto de zonificación que prohíba el desarrollo de la tierra que no sea de uso residencial. Supongamos que se invoca tal decreto para impedir que un terrateniente cree una porqueriza en su tierra, cuando la tierra de los vecinos se usa exclusivamente para fines residenciales. No podemos clasificar el decreto como una infracción de los derechos de propiedad del criador de cerdos sin decidir previamente que tales derechos incluyen el de infligir un daño a los vecinos por medio de una porqueriza. Y no podemos decidir esta cuestión previa sin evaluar los usos rivales afectados por el decreto. Sin embargo, una vez hecha tal evaluación, y una vez asignado consiguientemente el derecho de propiedad, no se sirve a ninguna función económica obligando a quienes ganan con el decreto a que compensen a los perdedores.[33]

El caso de la zonificación se encuentra entre la expropiación habitual de la tierra para un uso público, en el que debe pagarse una compensación justa, y la destrucción de un perjuicio, en el que no debe pagarse ninguna compensación (por la excelente razón de que si se pagara una compensación por la "confiscación" de un perjuicio se estimularía la creación de perjuicios). Consideremos el ejemplo relacionado de las leyes que facultan al gobierno a designar la fachada de un edificio como un monumento; tras la designación, el propietario no podrá modificar la fachada. Una alternativa para la designación sería la compra (posiblemente respaldada por la amenaza de condenación sujeta al pago de una compensación justa) por el gobierno de un derecho de servidumbre o permiso en lo tocante a la fachada. ¿Designa el gobierno "demasiados" monumentos cuando se sigue la ruta de la designación antes que la ruta de la compra? Esto es difícil de saber. El hecho mismo de que no haya ninguna compensación hará que los designados se resistan. La resistencia de los contribuyentes podría ser, en efecto, menor. Los programas de impuestos y gastos gubernamentales (los subsidios agrícolas, por ejemplo) son a menudo tan socialmente costosos como los programas reguladores, pero los costos se reparten tan ralamente entre el público contribuyente que pocos contribuyentes protestan.

¿Pero podría el gobierno designar los monumentos "errados", es decir, propiedades que valieran mucho más en un estado alterado? Posiblemente no, ya que entre mayor sea el valor alternativo mayor sería la resistencia a la designación. Sin embargo, existe el peligro de disminuir la oferta de monumentos bajo el enfoque de la designación: los dueños de edificios podrían apresurarse a demoler las fachadas de monumentos potenciales en previsión de la designación. De igual modo, el costo social de la ruta de designación, por oposición al costo financiero para el gobierno, podría ser mayor que si se

[33] Véase un análisis económico más detallado de las confiscaciones regulatorias en Thomas J. Miceli y Kathleen Segerson, "Regulatory Takings: When Should Compensation Be Paid?", 23 *J. Legal Stud.*, 749 (1994).

pagara una compensación debido a que los gastos de oponerse a la designación son costos sociales reales, y no sólo una transferencia de pagos.

Una última pregunta es: ¿Qué tan bien funciona en la práctica la compensación justa? Un estudio empírico del programa de renovación urbana de Chicago descubrió que, bajo el dominio eminente, las parcelas de valor elevado reciben sistemáticamente más que el valor de mercado justo, mientras que las parcelas de valor bajo reciben menos.[34] Tres razones explican este patrón. Primera, la capacidad del gobierno para variar sus insumos de servicios legales —de acuerdo con el valor de la parcela que habrá de confiscarse— se ve severamente limitada por las reglamentaciones que gobiernan la prosecución de un caso de confiscación. El resultado es una tendencia a que el gobierno gaste demasiado en la tramitación de un caso que involucre a una parcela de valor bajo, y demasiado poco en la tramitación de un caso que implique a una parcela de valor alto. Segunda, los costos fijos o de umbral de ir a juicio representan una porción mayor de los costos totales entre menos valiosa sea la parcela, y su efecto al estimular a los demandantes a arreglar tales casos a bajo costo no se contrarresta por completo por el incentivo que tiene el gobierno para formular ofrecimientos de arreglos generosos a fin de evitar los costos fijos de la litigación, porque éste puede ser capaz de dispersar estos costos entre cierto número de parcelas que está tratando de adquirir simultáneamente. Tercera, puede haber economías de escala adicionales para el gobierno cuando son homogéneas las parcelas que está adquiriendo simultáneamente. Esto permite la consolidación efectiva de los esfuerzos legales del gobierno: empíricamente, las parcelas de valor bajo tienden a ser más homogéneas que las parcelas de valor alto.

III.7. Perjuicio

La chimenea de una fábrica despide humos que ennegrecen la ropa y las telas de un área residencial cercana e incrementan la incidencia de enfermedades respiratorias. Como en el ejemplo de las chispas, el reto para el sistema jurídico consiste en asignar los derechos y las responsabilidades de tal modo que se minimice la suma de los costos del daño causado por el humo y se evite tal daño. Entre los ajustes posibles, la fábrica podría instalar equipo de supresión de humo, podría cerrar, o los casatenientes afectados podrían instalar equipos de limpieza del aire o mudarse más lejos de la fábrica. La determinación de cuál de estos métodos u otros sea más barato para la solución del conflicto de los usos de la tierra resulta más difícil que en el caso de las chispas, sobre todo porque los efectos de la contaminación sobre la salud humana no se entienden claramente y los costos estéticos de la contaminación

[34] Patricia Munch, "An Economic Analysis of Eminent Domain", 84 *J. Pol. Econ*., 473 (1976).

son difíciles de medir.[35] Y es crítica la elección de la correcta asignación inicial de los derechos; es probable que los elevados costos de transacción imposibiliten la corrección de una asignación inicial errada mediante subsecuentes transacciones de mercado.

Ha llegado el momento de examinar más de cerca las fuentes de los costos de transacción elevados. Un factor destacado por los economistas, que se menciona antes en este capítulo, es un gran número de partes de una transacción, aunque otros factores, tales como la incapacidad mental, figuran de forma importante en contextos legales particulares (véase más adelante IV.9). Otro factor que destacan los economistas es el monopolio bilateral: en su forma más simple, un mercado con sólo un vendedor y un comprador, ilustrado por el ejemplo del derecho de vía o por las negociaciones para arreglar una demanda.[36] Debido a que el demandante sólo puede arreglarse con el demandado, y el demandado sólo puede hacerlo con el demandante, hay una gama de arreglos dentro de los cuales cada parte preferirá el arreglo a la alternativa más costosa de la litigación. Mientras más grande sea la gama de precios, mayor será el tiempo y los recursos que consuman las partes en la negociación. En efecto, ambas partes pueden estar tan decididas a embolsarse la mayor parte de los beneficios potenciales de la transacción que nunca lograrán ponerse de acuerdo. Ambas partes pueden ser reacias a iniciar la discusión del arreglo, por temor a que una de ellas sugiera a la otra una buena disposición para llegar a un arreglo, ocasionando que la otra revalore a la baja el precio de reserva de la primera parte. (Aun así, los psicólogos apuntarían a una ventaja si se es el primero en hacer una oferta: tal ofrecimiento puede actuar como "punto de anclaje" en la negociación posterior, y así se dará lugar a un arreglo más cercano a éste que si no hubiera sido la primera oferta.)

Los costos de transacción alcanzan su mayor nivel cuando los elementos del monopolio bilateral coinciden con gran número de partes en la transacción. Por ejemplo, si los casatenientes tienen el derecho de verse libres de la contaminación, la fábrica que desee tener un derecho de contaminar deberá adquirir tal derecho de cada uno de los casatenientes. Pues si sólo uno de 1 000 rehúsa llegar a un arreglo, los derechos que la fábrica ha comprado a los otros 999 no valen nada. En virtud de que el que resiste (o el retenedor, como lo llamamos anteriormente: al retrasar el proyecto, "retiene" a la fábrica) puede sacar un precio exorbitante, como en nuestro ejemplo del derecho de vía de la sección anterior, cada uno de los casatenientes tiene un incentivo para

[35] Si el nivel de la contaminación varía geográficamente, el costo podrá estimarse comparando los valores de las propiedades, manteniendo constantes otros factores que pudieran afectar a estos valores. Véanse, por ejemplo, Timothy A. Deyak y V. Kerry Smith, "Residential Property Values and Air Pollution: Some New Evidence", 14 *Q. Rev. Econ. & Bus.,* 93 (1974); K. F. Wieand, "Air Pollution and Property Values: A Study of the St. Louis Area", 13 *J. Regional Sci.,* 91 (1973).

[36] Se discute con detalle después en XXII.5.

demorar su arreglo con la fábrica, con la esperanza de ser el único resistidor. En consecuencia, el proceso de negociación podrá prolongarse indefinidamente.[37]

Si en lugar de que los casatenientes tengan el derecho de verse libres de la contaminación la fábrica tiene el derecho de contaminar, aquéllos deberán juntarse y comprar el derecho de la fábrica si desean verse libres de la contaminación. Los costos de transacción serán de nuevo elevados. Cada uno de los casatenientes tendrá otra vez un incentivo para arrastrar los pies en sus negociaciones con la fábrica. Se dirá a sí mismo: "Si me niego a contribuir con mi parte justa del precio de compra, otros, que se preocupan más profundamente por la contaminación que yo, aportarán la diferencia. La fábrica se verá inducida a dejar de contaminar. Yo me beneficiaré junto con los demás, pero sin costo alguno". Los costos de la superación de esta resistencia mediante negociaciones entre los casatenientes afectados serán elevados si hay muchos de ellos.

Cuando hay costos de transacción considerables, la creación de derechos absolutos (es decir, sin calificaciones), para contaminar o para verse libre de contaminación, tenderá a ser ineficiente. Si los costos de transacción son prohibitivos, la fábrica que tiene el derecho absoluto de contaminar no tendrá ningún incentivo para cesar (o reducir) la contaminación, aunque el costo de la cesación fuese mucho menor que el costo de la contaminación para los casatenientes. A la inversa, los casatenientes que cuenten con un derecho absoluto de verse libres de la contaminación, si los costos de transacción son prohibitivos, no tendrán ningún incentivo para tomar sus propias medidas a fin de reducir los efectos de la contaminación, aunque el costo para ellos de hacer eso (quizá mudándose) sea menor que el costo para la fábrica de no contaminar o contaminar menos.

La alternativa para los derechos absolutos es equilibrar, que es el enfoque utilizado por el remedio más importante del derecho común para la contaminación: un "perjuicio", el cuasidelito de la interferencia irrazonable con el uso o el disfrute de la tierra.[38] El criterio de la racionalidad implica una comparación entre el costo para el contaminador de abatir la contaminación y el menor de los costos, para la víctima, de tolerar la contaminación o eliminarla ella misma. Éste es el estándar eficiente, pero nadie cree que el derecho de los perjuicios haya tenido jamás un gran impacto sobre la contaminación. La ausencia de contaminación es un bien superior, lo que quiere decir que se de-

[37] Esto sería algo menos probable en el caso del derecho de vía, incluso sin un dominio eminente, en virtud de que habría rutas alternativas.

[38] El allanamiento, el cual discutimos en la sección anterior, difiere del perjuicio porque es un ingreso no permitido a la tierra de alguien, por oposición a una interferencia con su uso o disfrute, pero ésta es una diferencia nominal antes que real. Raras veces (aunque sí ocurre) habrá una interferencia con el uso o disfrute de la tierra que no implique el ingreso al territorio de las ondas o partículas de alguna clase.

manda proporcionalmente más a medida que aumenta el ingreso, de modo que no es sorprendente que la demanda de abatimiento de la contaminación haya sido pequeña hasta hace poco tiempo, como lo es todavía en los países pobres. Más importante aún, a menudo tanto los contaminadores como las víctimas de la contaminación son numerosos, al igual que pequeños y difíciles de identificar, mientras que los daños médicos, estéticos y de otra clase son difíciles de medir. Por ejemplo, para abatir la polución de los automóviles se requerirían acciones de clase demasiado masivas, que implicaran a grupos enormes del lado de los demandantes y de los demandados en un litigio y problemas horribles de medición y de remedio. No es de extrañar que la promulgación de extensas regulaciones estatutarias de la contaminación (examinadas en el capítulo XIV) haya desplazado al remedio del perjuicio en las áreas principales de la controversia sobre la contaminación.

Supongamos que el perjuicio es anterior a la llegada de las víctimas que se quejan. Es posible que una fábrica se encuentre en un barrio que gradualmente se ha vuelto residencial, y que la contaminación generada por la fábrica disminuya el valor de la propiedad residencial en mayor medida que el costo de la cesación de operaciones de la fábrica. Bajo la doctrina llamada con justicia "ir hacia el perjuicio", los recién llegados no podrían demandar el cierre de la fábrica. Pero la mayoría de los estados rechazan esa doctrina, y tienen razón en hacerlo porque es posible que los costos de transacción impidan que el mercado se ajuste cómodamente al cambio de los valores relativos del uso industrial y residencial. ¿Debiera considerarse esto "injusto" para el dueño de la fábrica? No necesariamente; es posible que el precio que pagó por la tierra en primer lugar haya sido descontado para reflejar la posibilidad de que la fábrica tenga que cerrar algún día por los perjuicios que causa (véase después III.14). Y es probable que la reubicación no sea muy costosa para él.

Por supuesto, como se señala a menudo en este libro, siempre existe un riesgo de error cuando un tribunal trata de determinar los valores de mercado. En un caso de perjuicios, el tribunal hizo un esfuerzo ingenioso para minimizar este riesgo.[39] El comedero para el ganado de los demandados emitía olores que reducían el valor de un desarrollo residencial cercano, construido después de que el comedero había iniciado sus operaciones. El tribunal, en la demanda del desarrollador, ordenó que el comedero cerrara sus puertas como un perjuicio, pero con la condición de que el demandante pagara el que resultara menor de los costos del cierre o de la reubicación del comedero. Si, sabiendo que ése sería el fallo, el desarrollador hubiese previsto que podría ubicar el desarrollo en otra parte a un costo menor que el de la mudanza o el cierre del comedero, se habría minimizado el costo de los usos conflictivos de la tierra. Sin embargo, ésta no es una solución perfecta, porque un derecho a los costos del cierre o de la reubicación reducirá el incentivo de los dueños

[39] Spur Industries, Inc. *vs.* Del E. Webb Development Co., 494 P.2d 700 (Ariz., 1972).

de comederos para ubicarlos en puntos óptimos en relación con el desarrollo proyectado del área circundante.

Una dificultad paralela daña a otro enfoque del derecho común sobre el problema de la contaminación, ilustrada por el tratamiento legal del ruido de los aviones. Los dueños de aviones que vuelan a muy bajas alturas son responsables ante los dueños de las propiedades situadas un poco debajo de la ruta de vuelo por la disminución del valor de mercado generada por el ruido de los aviones, independientemente de que esa disminución supere o no las ganancias de la compañía aérea. Sin embargo, el dueño de la propiedad no puede denunciar la invasión como un allanamiento y, por lo tanto, no puede obligar a la línea aérea a negociar con él, pero sí puede obligarla a anular un derecho de servidumbre o permiso de ingreso para continuar sus sobrevuelos. Si el costo de los procedimientos de abatimiento del ruido es mayor que el daño del ruido para los propietarios, la aerolínea anulará dicho permiso. Si es menor, adoptará los procedimientos. Es concebible que, si el método más barato para el abatimiento del ruido resulta ser el aislamiento contra el ruido de las casas subyacentes, la aerolínea pagará tal aislamiento, ya que por hipótesis ese costo sería menor que la responsabilidad de la aerolínea por el daño causado por el ruido. Pero los altos costos de transacción podrían impedir este resultado, y en tal caso el enfoque del dominio eminente podría ser menos eficiente que un enfoque de perjuicios, por el que los costos del abatimiento del ruido para la aerolínea serían contrastados con los costos para los casatenientes del aislamiento contra el ruido, y si estos últimos costos fueran menores la aerolínea no sería responsable. Pero el enfoque del dominio eminente es mejor que un enfoque de allanamiento. Si los dueños subyacentes tuvieran derechos de propiedad que no pudieran ser obligados a vender contra el ruido causado por los aviones —en otras palabras, si pudieran lograr la prohibición de los sobrevuelos—, entonces cada propietario tendría un incentivo para hacerse el resistidor. La aerolínea, incapaz de comprar a un precio razonable todos los derechos de los dueños subyacentes a verse libres del ruido, tendría que cesar sus vuelos o adoptar procedimientos de abatimiento del ruido, incluso si la solución eficiente fuese que continuara generando ruido y los propietarios subyacentes lo toleraran o aislaran sus casas contra el ruido.

Sin embargo, un problema del enfoque del dominio eminente es que, una vez que la aerolínea concluye que los costos de los procedimientos de abatimiento del ruido son mayores que los beneficios de la disminución de la responsabilidad con los dueños subyacentes y, por lo tanto, adquiere derechos de servidumbre o de ingreso que autorizan un alto nivel de ruido, no tendrá ningún incentivo para reconsiderar la adopción de tales procedimientos cuando sus costos bajen o su eficacia aumente, porque el beneficio de un nivel de ruido menor en el futuro sería obtenido enteramente por los dueños subyacentes. Este problema podría resolverse creando derechos de servidum-

bre o permisos de ingreso del ruido por un tiempo limitado.[40] Pero la solución crearía un problema nuevo. Los dueños de inmuebles que otorguen derechos de servidumbre o permisos de ingreso de ruido perpetuos tendrán de aquí en adelante todos los incentivos para la adopción de cualquier medida de disminución del ruido que cueste menos de lo que aumenta el valor de su propiedad. No tienen este incentivo bajo un sistema de derechos de servidumbre o permisos de ingreso de tiempo limitado, porque cualquier medida tomada por un dueño de inmuebles que disminuya el daño del ruido disminuirá en una cantidad igual el precio que reciba por el permiso de ingreso del ruido en el periodo siguiente.

III.8. Soluciones para el problema de los usos incompatibles de la tierra mediante convenios de fusión y restricción

La obtención de la solución eficiente en el caso de las chispas, el caso de los humos de la fábrica y nuestros otros ejemplos de usos conflictivos de la tierra habría sido mucho más simple si un solo individuo o empresa hubiera sido el único propietario de toda la tierra afectada. Un solo dueño de la fábrica y de la propiedad residencial afectada por el humo desearía maximizar el valor combinado de ambas propiedades. Ésta es la meta económica correcta, y el esfuerzo por alcanzarla no se vería obstruido por los costos de la obtención del acuerdo de muchos propietarios separados.

¿Por qué son, entonces, tan poco frecuentes tales fusiones? En primer lugar, la compra de todas las propiedades afectadas por alguna fuente de contaminación implicaría costos de transacción onerosos porque requeriría una transacción con muchos propietarios, por lo que habría el mismo tipo de problema de retención que se discutió anteriormente en el contexto del dominio eminente. En segundo lugar, es posible que una sola empresa no sea capaz de operar eficientemente en mercados no relacionados —producción fabril e inmuebles residenciales, ferrocarril y agricultura, administración aeroportuaria y administración inmobiliaria—.[41] El solo tamaño podría ser también una fuente de costos, debido a la presencia tanto de deseconomías como de economías de escala.

[40] Como lo propusieron William F. Baxter y Lillian R. Altree en "Legal Aspects of Airport Noise", 15 *J. Law & Econ*., 1 (1972). Si la servidumbre o permiso de ingreso se limita (digamos) a 10 años, la aerolínea revisará periódicamente el avance de la tecnología del abatimiento del ruido a fin de determinar si la adopción de procedimientos del abatimiento en cuestión le ahorraría más dinero disminuyendo sus costos esperados de la servidumbre en mayor medida que la suma que los procedimientos añadirían a su capital y sus costos de operación.

[41] Es quizá posible que se contrarresten algunos de los costos de la subespecialización mediante el arrendamiento, pero la coordinación de los arrendatarios podría ser casi tan costosa como el mercado, como veremos más adelante. La pérdida de control en las organizaciones se discute en la sección siguiente de este capítulo y también después en XV.1.

A la solución de la propiedad única se aproximan aquellos estados productores de petróleo y gas (la mayoría) en los que es posible la unificación obligatoria, por la que el voto de una mayoría considerable (de ordinario dos tercios) de los propietarios de un campo de petróleo o de gas para que se opere el campo como si fuese de propiedad común será vinculante para la minoría. (De nuevo, un requerimiento de unanimidad sería ineficiente porque crearía un problema de retención.) Los propietarios de los derechos de petróleo y de gas estarán bombeando desde un yacimiento subterráneo de propiedad común. Por lo tanto, cada uno de tales propietarios tendrá un incentivo para perforar muchos pozos a fin de bombear la mayor cantidad posible de petróleo o de gas con la mayor celeridad posible, aunque el costo total de la producción del campo en conjunto pudiera disminuir, y el rendimiento total de petróleo o de gas pudiera incrementarse, perforando menos pozos y agotando el recurso más lentamente. Cualquier cantidad de petróleo o gas que no bombee el perforador de un yacimiento subterráneo la bombeará otro perforador.

Una solución alternativa al problema de los usos incompatibles de la tierra es el convenio de restricción. El desarrollador de un terreno que desea maximizar el valor de la propiedad pero no administrarla puede incluir restricciones en cada escritura de venta contra los usos de la tierra que reducirían el valor de la propiedad en su conjunto. Tales restricciones "van de la mano con la tierra", lo que quiere decir que son aplicables (si están consignadas en un registro de títulos de tierra) a cualquier propietario futuro de la tierra, así como al propietario actual, y son aplicables por los sucesores de los compradores originales. Una mera obligación contractual para el comprador y sus sucesores *ante el desarrollador* no sería suficiente. Una vez que éste completara el desarrollo del terreno, tendría escaso o ningún interés en hacer efectiva la obligación. De todos modos, tal contrato no sería vinculante para un comprador del comprador original; el segundo comprador quizá ni siquiera podría saber de la existencia del contrato que su predecesor hubiese celebrado con los vecinos. Además, no se puede estar seguro de que el segundo comprador acate la restricción voluntariamente, porque, aunque por hipótesis la restricción aumenta el valor de todo el terreno, si todos los demás acatan la restricción entonces la tierra del propietario que no la acate será más valiosa aún; recibirá los beneficios de la restricción sin incurrir en alguno de los gastos. Así que el comprador acatará la restricción sólo si se le paga para que lo haga. Pero no se le pagará… o, por lo menos, no mucho. Si los otros propietarios son numerosos, tendrán dificultades para superar el problema de los resistidores en su seno. Y ganarán poco si lo hacen, porque si el propietario actual vende, los demás propietarios tendrían que transar de nuevo con el comprador.

Existen limitaciones respecto al tipo de promesa (a menudo llamada *derecho de servidumbre* o *permiso de ingreso*) que acompaña a la tierra (es

decir, que es aplicable como un derecho de propiedad más que como un simple derecho de contrato). Supongamos que el tenedor de un terreno vende una parte de su tierra y le promete al comprador que no venderá bienes o servicios en competencia con él. ¿Acompañará esta promesa a la tierra del vendedor? El derecho común contesta que no; la promesa no "afecta e involucra" a la tierra. (Otra palabra para los permisos de ingreso que acompañan a la tierra es servidumbre "anexa", servidumbres que se refieren a la tierra.) Sin embargo, si el vendedor prometió no construir una cerca que obstruyera la vista del comprador en la tierra que retiene, esta promesa acompañaría a la tierra, es decir, sería aplicable contra los sucesores del vendedor incluso si ellos no renovaran la promesa, debido a que la promesa involucra un uso real de la tierra.

Una razón para la distinción es el impedimento a la asignación eficiente de recursos que se crearía al aplicar restricciones anticuadas sobre el uso de la tierra. Otra es que es difícil llevar un registro de las promesas que no se ocupan del propio uso de la tierra (la tierra retenida del vendedor en los ejemplos anteriores), aunque este problema se mitiga en gran medida a través de un sistema de registro público de intereses de la tierra. De este modo, no es de sorprender que hasta que se creó en Inglaterra un registro de la titularidad de la tierra, en 1925, los tribunales ingleses sólo aplicaran permisos de ingreso o derechos de servidumbre negativos, como en el caso de la servidumbre de vista (mencionado antes), el apoyo a la construcción (por ejemplo, el prestado por las casas contiguas a una casa en una hilera), los derechos de acceso (si la propiedad estaba rodeada por otra propiedad) y los derechos ribereños (la igualdad de derechos de los propietarios de tierras ubicadas a la orilla de un río u otro cuerpo de agua para hacer uso de él). Las violaciones de tales permisos de ingreso o derechos de servidumbre pueden observarse fácilmente. En cambio, si el vendedor que promete al comprador que no ofertará bienes o servicios en competencia con los de éste vende su tierra, ¿cómo podría determinar su comprador si al adquirir las tierras ha adquirido obligaciones, que no tienen relación con su uso de la tierra, con algún vecino?

Por lo tanto, una manera de ver el requerimiento de "afecta e involucra [a la tierra]" es que aplica los derechos que aumentan el valor de una propiedad, independientemente de quién la posee.[42] Si el valor de una promesa de no competencia es beneficioso o no para el promisario depende del uso particular que haga de su tierra (en el ejemplo, la venta de artículos que el

[42] Véase Antony Dnes y Dean Lueck, "Asymmetric Information and the Law of Servitudes Governing Land", 38 *J. Legal Stud.*, 89, 98 (2009). Para comparar puntos de vista incompatibles sobre la conveniencia del requerimiento de "afecta e involucra [a la tierra]" compárese a Richard A. Epstein, "Notice and Freedom of Contract in the Law of Servitudes", 55 *S. Cal. L. Rev.*, 1353 (1982), con Jeffrey E. Stake, "Toward an Economic Understanding of Touch and Concern", 1988 *Duke L. J.*, 925.

promitente hubiera querido vender para competir con él), y no existe la suposición de que la promesa aumentaría el valor de esa tierra para cualquier otro propietario.

Esta discusión destaca la diferencia económica existente entre los derechos de propiedad y los contractuales. Un derecho de propiedad excluye (en el límite) a todo el resto del mundo del uso de una cosa, excepto bajo los términos del dueño. Un derecho contractual excluye sólo a la otra parte del contrato. La libertad para contratar y aplicar contratos, pero no para crear derechos de propiedad legalmente exigibles, no optimizaría el uso de los recursos. Si A le compra a B el derecho de trabajar la tierra de B, pero B no tiene el derecho de excluir a otros de ese trabajo, A (como B antes de él) no tendrá ningún incentivo para explotar la tierra óptimamente. De igual modo, sin derechos de propiedad, el problema del pastoreo excesivo en nuestro ejemplo del pasto común no se resolvería, aunque los agricultores que usan el pasto vendieran sus derechos a un solo individuo o empresa. Después de que el nuevo propietario hubiera reducido el congestionamiento cobrando a los agricultores que le hubieran vendido sus derechos una cuota apropiada por la continuación del uso del pasto, otros agricultores empezarían a pastar *sus* propias vacas en el pasto: no tendrían ninguna obligación de pagar una cuota.

Como solución para los usos de tierra incompatibles, los convenios restrictivos tienen dos limitaciones. Primero, generalmente son viables sólo en el contexto especial de la propiedad única inicial de un área grande. No proveen ninguna solución para el problema de contaminación típico, pues es raro que un área suficientemente grande para albergar una fábrica y a todas o la mayoría de las residencias afectadas por su humo se encuentre bajo una propiedad común. En segundo lugar, tales convenios son inflexibles frente a los cambios que pueden modificar los valores relativos de los usos conflictivos de la tierra. El propietario que desea dar a su tierra un uso prohibido por un convenio restrictivo deberá obtener el consentimiento de todos los dueños de propiedades en cuyo favor opera el convenio; si éstos son muchos, los costos de transacción pueden ser prohibitivos. Por lo tanto, algunos convenios establecen que expirarán después de cierto número de años a menos que sean renovados por el voto mayoritario de los terratenientes afectados. Y los tribunales podrían negarse a aplicar un convenio restrictivo alegando que es obsoleto; que el uso prohibido de la tierra es ahora claramente más valioso que el uso protegido por el convenio. Esto es como la doctrina *Cy pres* mediante la que se faculta a los tribunales para dilucidar los términos de fideicomisos de beneficencia, según veremos en el capítulo XIX.

El problema del convenio obsoleto sería menos grave si los tribunales se negaran a sancionar los incumplimientos de los convenios restrictivos y en su lugar se limitaran a condenar al pago de daños a los demandantes victoriosos. La responsabilidad del pago de daños no evitaría un incumplimiento

que aumentara el valor de la propiedad del demandado en mayor medida de lo que disminuye el valor de las demás propiedades del predio, ya que, por hipótesis, su responsabilidad sería menor que la ganancia derivada del incumplimiento. En cambio, una orden judicial coloca al violador potencial en la misma posición que la de la aerolínea, que es condenable por los dueños de las propiedades subyacentes, o la del ferrocarril, que es condenable por allanar la propiedad que requiere para completar su derecho de vía. Para lograr que se levante la prohibición, el violador potencial tendrá que negociar con cada uno de los derechohabientes, quizá deba pagar un precio exorbitante a unos cuantos resistidores y quizá no logre completar siquiera la transacción.

La inflexibilidad de los convenios restrictivos ha llevado a un número cada vez mayor de desarrolladores a establecer asociaciones de casatenientes facultados para modificar las restricciones existentes a los usos que pueden dar a su propiedad. Este método para la solución del problema de los costos de transacción elevados se asemeja a otro método: el de la empresa comercial.[43] Además de estas soluciones privadas para el problema de los usos conflictivos de la tierra, existe una solución pública: la zonificación. Deben distinguirse dos tipos de ella: la de separación de los usos divide a una ciudad u otra unidad gubernamental local en zonas, y permite sólo ciertos usos de la tierra en cada una, de modo que podría haber zonas separadas para altos edificios de departamentos, para casas unifamiliares, para negocios, para fábricas, etc. La zonificación excluyente (un término empleado a menudo peyorativamente, pero que aquí se emplea de manera neutral), adoptada de ordinario por unidades más pequeñas que una ciudad o un condado, trata de excluir por completo ciertos usos de la tierra, más que sólo segregarlos; un suburbio que requiere lotes de tamaño mínimo estaría practicando la zonificación excluyente. La interrogante principal acerca de la zonificación por separación de los usos consiste en saber si produce tanta diferencia,[44] especialmente teniendo en cuenta la alternativa de convenios restrictivos.[45] Es improbable que encontremos una casa y una fábrica contiguas aunque no haya ninguna zonificación. Los inmuebles residenciales tienen de ordinario un precio mayor que los empleados para fines industriales, ya que un terreno mucho más grande es más adecuado para uso industrial que para uso residencial, de modo que el dueño de una fábrica no desearía construirla en un área residencial a menos que su propósito fuese la extorsión, en cuyo caso el derecho de los perjuicios podría resolver el problema con eficacia.

La zonificación excluyente tiene más probabilidades de afectar el uso de la tierra que la zonificación por separación de los usos. Un lote grande valdría tal vez más si se empleara para la construcción de un edificio elevado de

[43] Véase el capítulo XV. Otra solución es la tributación. Véase después en XIV.5.

[44] Véase una respuesta negativa en Bernard H. Siegan, *Land Use without Zoning,* 75 (1972).

[45] Véase Christopher Berry, "Land Use Regulation and Residential Segregation: Does Zoning Matter?", 3 *Am. Law & Econ. Rev.*, 251 (2001).

departamentos que si se utilizara para la construcción de una sola casa, por lo menos si se omiten los efectos sobre otros casatenientes de la comunidad, como lo haría de ordinario el desarrollador. Estos efectos podrían incluir el congestionamiento de carreteras y estacionamientos, y las cargas para los servicios municipales como las escuelas públicas. Sin embargo, adviértase que si los residentes del edificio elevado deben pagar los costos adicionales que imponen por el uso de las escuelas públicas y las calles, no habrá ninguna externalidad que justifique la zonificación excluyente. Aunque ésta podría ser eficiente en principio, en la práctica podría ser muy diferente porque el incentivo de los funcionarios —las personas que redactan ordenamientos de zonificación y los aplican— podría alejarlos de la meta de la eficiencia, como podremos ver en los capítulos XX y XXIV. Además, la zonificación excluyente tiende a redistribuir la riqueza de los pobres a los ricos. Ninguna comunidad es perfectamente homogénea; por consiguiente, la gente pobre que viva en una comunidad de ricos se beneficiará de los servicios públicos con los impuestos que paga la gente rica… a menos que la zonificación excluyente haga que vivir en esa comunidad sea demasiado caro para la gente pobre.

III.9. Prohibiciones en comparación con el pago de daños como métodos de aplicación de los derechos de propiedad

Varias veces hemos visto que se distinguen los derechos de propiedad de otras clases de intereses legales en términos de los remedios, y ahora podemos generalizar ese punto: en las situaciones de usos conflictivos en las que los costos de transacción son bajos, como en un caso común de allanamiento, debería permitirse de ordinario el remedio de la prohibición; al obligar al demandado (quien violó el derecho del demandante) a comprar el derecho en caso de que sea más valioso para él, facilita la determinación por el mercado del valor. Pero cuando los costos de transacción son altos, impidiendo tal transacción correctiva, se facilita la asignación de los recursos a sus usos más valiosos negando a los dueños de la propiedad un remedio de prohibición y limitando en cambio al demandante a un remedio de pago de daños que iguale su estimación del costo de la violación de sus derechos, con lo que se permite que continúe la violación si para el invasor vale más que lo que le cuesta a la víctima. De este modo, el régimen del allanamiento, al negarse a considerar el valor de la actividad del invasor, canaliza la transacción hacia el mercado, al que pertenece, mientras que el régimen de los perjuicios, al permitir una comparación del valor de las actividades en conflicto, simula el resultado de una transacción de mercado que probablemente no es factible debido a los costos de transacción.[46]

[46] Véanse algunos argumentos en el sentido de que la ley sigue generalmente este patrón en

Sin embargo, este análisis es sensible a la cantidad y la corrección de la información poseída por los tribunales.[47] Podría parecer que, si éstos tienen el conocimiento exacto de los valores de los usos en conflicto, no habrá nada que escoger entre los derechos de propiedad aplicados por prohibiciones y las reglas de responsabilidad aplicadas por los fallos de pagos de daños. De cualquier modo, el tribunal será capaz de duplicar los resultados del mercado, siempre que la responsabilidad sea estricta, es decir, que la persona que toma el automóvil de usted de la cochera sea obligada a pagarle daños iguales a la cantidad total en la que usted valúa el automóvil, aunque esa persona lo valúe más que usted;[48] de otro modo, habrá gastos excesivos en la toma, y la prevención de la toma, de la propiedad. Pero, en realidad, bajo estos supuestos es superior el enfoque de la prohibición al enfoque de pago de daños, ya que la adquisición de información es costosa y se la requiere menos para determinar si el uso del demandante es más valioso que el uso del demandado que para determinar el valor del uso del demandante (o la disminución de tal valor por obra del demandado), así como determinar si una persona es más alta o pesada que otra es más fácil que determinar cuán alta y pesada es cualquiera de esas dos personas.

Sin embargo, sólo raras veces puede un tribunal determinar correctamente los valores de los usos rivales, sobre todo cuando están implicados valores subjetivos. Podría parecer que este supuesto hace girar decisivamente la elección en favor del pago de daños. Mientras no sea más probable que el tribunal sobrestime los daños en lugar de subestimarlos, el otorgamiento típico de daños será una aproximación razonable al daño promedio sufrido por la propiedad del demandante, de modo que creará los incentivos correctos. Dado que un tribunal no será capaz de estimar el valor del uso del demandante con mayor apego que el valor del uso del demandado, el enfoque del pago de daños tiene sobre el enfoque de la prohibición la ventaja de no desalentar el uso por parte del demandado si es más valioso que el uso del demandante. En tal caso, el demandado pagará daños al demandante, pero no cesará su propio uso si obtiene un valor mayor que el costo del fallo. El enfoque de los daños permite que el demandado utilice su conocimiento privado de los costos de los daños y de la prevención para contrarrestar las consecuencias ineficientes de la falta de conocimiento de estas cosas por parte del tribunal.

Comoquiera que sea, no otorgar nunca el remedio de la prohibición sería ineficiente, porque entonces los tribunales tendrían que determinar los daños

William M. Landes y Richard A. Posner, *The Economic Structure of Tort Law,* 42-44 (1987); Thomas W. Merrill, "Trespass, Nuisance, and the Costs of Determining Property Rights", 14 *J. Legal Stud.,* 13 (1985).

[47] Véase en términos generales Louis Kaplow y Steven Shavell, "Property Rules *versus* Liability Rules: An Economic Analysis", 109 *Harv. L. Rev.,* 713 (1996).

[48] En otras palabras, el sistema hipotético descrito aquí es un sistema de dominio eminente privado con una compensación plena antes que "justa".

sufridos por el demandante en cada caso, en lugar de sólo tener que determinar (cuando los costos de transacción son bajos) si tales daños exceden al valor del uso rival del demandado (una determinación más fácil de hacer). En un régimen de no prohibición surgiría otro problema. Supongamos que, debido a la dificultad que representa para los tribunales la determinación correcta del pago de daños, cada dueño de propiedades espera recibir un pago típico de daños por la toma ilegal u otra invasión de su propiedad; entonces quienes esperen que sus daños sean mayores que el promedio emplearán medidas precautorias, mientras que quienes esperen que sus daños estén por debajo del promedio no tomarán tal vez ni las precauciones mínimas, con la esperanza de que la propiedad sea tomada. Cuando están implicados valores subjetivos, es probable que los tribunales subestimen sistemáticamente los daños porque el demandante tiene la carga de la prueba y se desechan las pruebas muy especulativas.

Un último argumento a favor del remedio de la prohibición es que existe el peligro del reciclaje si se permite que unos tomen la propiedad de otros recíprocamente, sujetos al pago del valor determinado por el tribunal. Si A valúa su automóvil en 20 000 dólares, B lo toma y el tribunal ordena que B le pague a A 18 000 dólares, A tendrá un incentivo para recobrarlo de manos de B, esperando que el tribunal le permita conservar el automóvil por 18 000 dólares, lo que le rendirá una ganancia neta de 2 000 dólares. Los tribunales podrían tratar de prevenir la nueva determinación del valor aplicando doctrinas de preclusión (la cosa juzgada o el impedimento colateral; véase después XXII.14), pero eso sería difícil porque los valores cambian a lo largo del tiempo.

Si la relación de la prohibición con el caso del costo de transacción bajo y el pago de daños con el caso del costo de transacción elevado capta la ley y la economía de las disputas sobre infracciones de la propiedad, ¿cómo podemos explicar el famoso caso Boomer,[49] donde, en lugar de aplicar el estándar habitual del uso razonable, el tribunal sostuvo que un perjuicio es cualquier interferencia sustancial con el disfrute de su tierra por parte de otra persona? Ésta es un estándar típico del allanamiento, y había pocos demandantes (y un solo demandado: una fábrica de cemento que emitía polvos), de modo que podría haberse pensado que el tribunal otorgaría una prohibición, la que sólo podría haberse cumplido con el cierre de la planta. El tribunal se negó, sosteniendo que ejercería su discreción equitativa para no imponer la prohibición si el demandado compensaba a los demandantes por el daño que les había causado el perjuicio.

Éste era el fallo errado si los costos de transacción hubiesen sido bajos, pero es posible que hayan sido elevados aunque las partes fuesen pocas. Pues

[49] Boomer *vs.* Atlantic Cement Co., 257 N.E. 2d 870 (Nueva York, 1970). Véase en términos generales Jeff L. Lewin, "*Boomer* and the American Law of Nuisance: Past, Present, and Future", 54 *Albany L. Rev.,* 191 (1990).

este caso fue un ejemplo dramático del monopolio bilateral. La planta de cemento había costado 45 millones de dólares, y ésta es, al menos, una estimación burda de la pérdida si la planta se hubiese visto obligada a cerrar, aunque la pérdida efectiva podría haber sido mayor o menor. El costo del perjuicio para los demandantes era de sólo 185 000 dólares. Por lo tanto, *cualquier* precio que se pagara por la eliminación de la prohibición y que se encontrara entre 185 000 y 45 millones de dólares (si ése fuera el valor actual de la planta) habría dejado a ambas partes en una mejor situación que la que se crearía si se aplicara la prohibición. Ése es un intervalo de negociación *enorme*; a cada una de las partes le habría convenido invertir recursos sustanciales para recibir la mayor suma posible. Esta cara negociación se evitó con el novedoso enfoque remediador del tribunal, aunque una alternativa habría sido que, mediante un balanceo de los costos de las partes respectivas, se hubiera decidido que la planta no constituía un perjuicio.

El problema de los usos de la tierra incompatibles, que hemos venido analizando en términos de la definición y la transferencia de los derechos de propiedad, se aborda con frecuencia (como hemos empezado a hacerlo en esta sección) en términos de las "externalidades". El daño causado a los cultivos del agricultor por las chispas de la máquina es un costo del transporte ferroviario que el ferrocarril no tendrá en cuenta al tomar sus decisiones a menos que se vea obligado por la ley a hacerlo o que sea el dueño de la tierra agrícola; el costo es externo a su proceso de toma de decisiones.[50] El término es útil pero potencialmente engañoso. Sugiere que la solución correcta en el caso de las chispas consiste en imponer una responsabilidad al ferrocarril, aunque recordemos de la discusión anterior del caso de las chispas de la locomotora que en la teoría económica no hay ninguna presunción de que el ferrocarril, y no el agricultor, debiera pagar el costo del daño causado por las chispas. Aun si se define la "externalidad" como algo exterior a los procesos de decisión del mercado, antes que a los del victimario, sigue siendo un uso potencialmente engañoso, porque si los costos de transacción son bajos el mercado podría internalizar las externalidades.

III.10. La propiedad dividida: los derechos "divididos" sobre las tierras o bienes raíces

Es posible que más de una persona tenga un derecho de propiedad sobre la misma cosa. Nuestro recurso común fue un ejemplo; otro ejemplo más tradicional es el de los diferentes "patrimonios" en terrenos. Los derechos de propiedad de bienes raíces pueden dividirse entre un inquilino de por vida y un

[50] También existen beneficios externos. Es probable que la apertura de una tienda nueva en un barrio, al atraer nuevos compradores, cree clientes nuevos para las tiendas cercanas.

propietario del resto, entre inquilinos conjuntos (un tipo especial de la copropiedad), entre un arrendatario y un terrateniente, y en otras formas. Tales divisiones (ya sean concurrentes pero no excluyentes, o excluyentes pero de duración limitada) crean incentivos para el uso ineficiente similares a los que crea la propiedad separada del derecho de vía del ferrocarril y de la tierra agrícola adyacente, o del aeropuerto y la comunidad residencial adyacente.

Este problema se ha discutido extensamente en conexión con la pobreza de Irlanda en el siglo XIX.[51] La mayoría de los agricultores eran arrendatarios; y parecería que un inquilino tendría escaso incentivo para mejorar la tierra, porque todo mejoramiento que sobreviviera al periodo del arrendamiento conferiría al terrateniente un beneficio no compensado bajo la doctrina de los accesorios: todo lo que el arrendatario adhiera a la propiedad se convierte en propiedad del terrateniente a la expiración del arrendamiento. (De otro modo, el inquilino estaría imponiendo costos al terrateniente si a éste no le gustara el artículo adherido y removerlo fuera costoso.) Pero esta sugerencia parece violar el Teorema de Coase a primera vista. ¿Por qué no aceptaban los terratenientes, en los contratos de arrendamiento, compensar a los inquilinos por los mejoramientos, por ejemplo, otorgándoles un porcentaje de los ingresos netos derivados de la tierra después de la expiración del contrato?

Había, y hay todavía, acuerdos de repartición entre terratenientes y arrendatarios, en particular la aparcería, pero no siempre funcionan. Supongamos que el terrateniente acepta proveer la tierra, las semillas y los fertilizantes, y que el agricultor acepta proveer la mano de obra, repartiéndose por partes iguales los ingresos de la venta de las cosechas. Los resultados no serían óptimos, como veremos con un ejemplo sencillo. Supongamos que si el agricultor trabajara una hora extra cada semana en el mejoramiento de la tierra aumentaría en dos el valor en dólares de la producción de la granja (descontando todos los costos adicionales además de su tiempo), y que el costo de oportunidad —o precio sombra[52]— de su tiempo en términos del ocio sacrificado es de sólo 1.50 dólares. La eficiencia requiere que él trabaje la hora extra, pero él no lo va a hacer, porque según su trato con el terrateniente recibirá sólo 1 dólar por el trabajo que a él le costó 1.50 dólares. Se requerirá un acuerdo de repartición más complicado para obtener resultados óptimos, y entre más complicado sea, más costosos serán su negociación y su cumplimiento forzoso. Y este ejemplo no considera el problema de las mejoras a largo plazo, suponiendo que el inquilino estará allí todavía cuando las mejoras fructifiquen. Es posible que no se encuentre allí si el arrendamiento es a corto plazo.

El problema será menos grave entre más largo sea el contrato; por lo tanto, no es sorprendente que se desarrollara un sistema de derechos con-

[51] Véase, por ejemplo, A. C. Pigou, *The Economics of Welfare,* 174-175, 178-183 (4ª ed., 1932); Barbara Lewis Solow, *The Land Question and the Irish Economy, 1870-1903* (1971).

[52] En el capítulo I vimos que el precio sombra es el precio que pagaría una persona por algo que no se intercambia en ningún mercado.

suetudinarios del arrendatario que dificultaba el desalojo del inquilino por parte del terrateniente, en forma directa o indirecta (esto último aumentando las rentas hasta que el inquilino se viera obligado a renunciar al contrato). Todavía habría habido un problema de incentivos del arrendatario si fuese probable que las mejoras óptimas del inquilino sobrevivieran a los inquilinos, pero las grandes mejoras de capital, las del tipo que tenían mayores probabilidades de sobrevivir al arrendatario actual, tenían que ser hechas de todos modos por los terratenientes y no por los arrendatarios, porque aquéllos tenían el capital. Así que, si acaso, el problema no era que los arrendatarios carecieran de incentivos para mejorar la tierra, sino que los derechos consuetudinarios del arrendatario les hacían difícil a los terratenientes recuperar el costo de sus propias mejoras cobrando rentas mayores, ya que el inquilino podría quejarse de que el aumento de la renta violaba sus derechos consuetudinarios.

No hay una solución sencilla para el problema de la propiedad dividida, fuera de la propiedad única… pero ésta tampoco es una solución sencilla. Si el arrendatario es transformado en empleado del terrateniente, el problema de la propiedad dividida desaparece, pero se ve sustituido por un problema enteramente análogo de la negligencia del agente debido a que el empleado no se queda con cada dólar de la producción obtenida por su trabajo, justo como ocurre con el arrendatario. Y es posible que el arrendatario no esté dispuesto a comprarle la granja al terrateniente (a pesar de que esto eliminaría el problema de la negligencia), aunque pudiera alcanzarle para hacerlo, debido al riesgo adicional en el que incurriría (lo que ilustra la observación importante de que el arrendamiento es una forma de la dispersión del riesgo).[53]

Sería de esperarse que los tribunales interpretaran los contratos de arrendamiento como si la intención de las partes hubiese sido que la propiedad fuese administrada por el arrendatario como si fuese el propietario;[54] porque presumiblemente ésa *fue* la intención de las partes, si son maximizadores racionales de la ganancia. Esta política se refleja en la doctrina del derecho común del desperdicio, que media entre los intereses rivales de los inquilinos de por vida y los dueños del resto. Un inquilino de por vida tendrá un incentivo para maximizar no el valor de la propiedad, es decir, el valor presente de toda la corriente de ingresos futuros obtenibles de ella, sino sólo el valor presente de la corriente de ingresos obtenibles durante el ciclo esperado de su vida. Por lo tanto, el inquilino deseará cortar la madera antes de que haya alcanzado su crecimiento completo —aunque el valor presente de la madera sería mayor si se pospusiera la tala de una parte o la totalidad de los árboles—. El derecho del desperdicio prohíbe que los inquilinos de por vida modifiquen la propiedad de modo que se reduzca el valor de la propiedad en su conjunto.

[53] Véase Steven N. S. Cheung, *The Theory of Share Tenancy* (1969).

[54] Véase un ejemplo excelente en Suydam *vs.* Jackson, 54 N.Y. 450 (1873).

Éste es el punto de vista judicial moderno tradicional del derecho, pero se ha impugnado como inferior a la regla que definiría el desperdicio como cualquier "modificación sustancial" o "alteración material" de la propiedad por parte del inquilino de por vida; se dice que es más clara que el enfoque de la disminución del valor.[55] En realidad parece ser menos clara. Además de las situaciones en las que la "modificación" aumenta el valor de la propiedad en su conjunto, no todos los cambios en la condición de la propiedad podrían pensarse de manera convincente como una alteración *material*. ¿Es un desperdicio instalar una central de aire acondicionado? ¿Convertir un granero en un garaje? ¿Pavimentar un camino de entrada? ¿Todas las mejoras que modifican una propiedad son un desperdicio? Sin embargo, existe un buen argumento, aunque no necesariamente convincente en grado sumo, del enfoque de la alteración material, aunque sólo es aplicable para la propiedad residencial: el propietario podría tener gustos idiosincrásicos y sentirse peor como consecuencia de la modificación del inquilino a pesar de que ésta haya aumentado el valor de mercado de la propiedad.

Podría parecer que no había ninguna necesidad de un derecho del desperdicio, porque el inquilino de por vida y el dueño del resto negociarían un plan óptimo para la explotación de la propiedad. Pero en virtud de que el inquilino y el dueño del resto sólo pueden contratar el uno con el otro, la situación es de nuevo de monopolio bilateral, y los costos de transacción pueden ser elevados. Además, los dueños del resto pueden ser niños sin capacidad legal para celebrar contratos vinculantes; pueden ser incluso niños nonatos todavía. El problema del monopolio bilateral es menos agudo en el caso del terrateniente y el inquilino, porque los términos de un contrato de arrendamiento se establecen antes de que ambos queden atrapados en una relación recíproca. Pero con gran frecuencia se crea un arrendamiento de por vida mediante un testamento, y es posible que el legatario (para quien la planeación de la propiedad puede ser una experiencia única en la vida) no esté alerta ante los conflictos que pueden surgir entre los inquilinos a perpetuidad y los dueños del resto.

El derecho del desperdicio ha sido sustituido en gran medida por un método de administración de la propiedad más eficiente, semejante a la unificación: el fideicomiso. Colocando la propiedad en un fideicomiso, el otorgante puede dividir el interés del beneficiario cuantas veces quiera sin preocuparse por la división de la propiedad. El fiduciario administrará la propiedad como una unidad, maximizando su valor y repartiendo ese valor entre los beneficiarios del fideicomiso en las proporciones deseadas por el fideicomitente. Por supuesto, es necesario que el fiduciario reciba los incentivos apropiados para hacer esto.[56] Para este propósito se emplean lo mismo la táctica de estí-

[55] Thomas W. Merrill, "Melms *vs*. Pabst Brewing Co. and the Doctrine of Waste in American Property Law", 94 *Marquette L. Rev.*, 1055 (2011).

[56] Por lo que toca al problema de la compensación apropiada del fiduciario, véase W. Bishop y D. D. Prentice, "Some Legal and Economic Aspects of Fiduciary Remuneration", 46 *Mod. L.*

mulos como la de amenazas: tanto la zanahoria como el palo. Se compensa al fiduciario, pero también se le impone la obligación fiduciaria de administrar el fideicomiso como si fuera de su propiedad (y tuviera las mismas preferencias, incluida la actitud hacia el riesgo, que se sabe que tienen los beneficiarios o se puede esperar que éstos tengan).

El inquilino no tiene siempre un horizonte de tiempo más corto que el del dueño del título simple (*fee simple*), la propiedad absoluta (en el caso de la relación ordinaria entre el terrateniente y el arrendatario) o el del dueño del resto (en el caso de un arrendamiento perpetuo). Veamos el caso de un arrendamiento petrolero que promete al arrendador una regalía fija por barril. A menos que espere que el ritmo del aumento del precio del petróleo supere a la tasa de interés, el arrendador deseará que el petróleo sea extraído con la mayor celeridad posible, independientemente de que el campo sea unificado o no. Esto significará la perforación de demasiados pozos. Pero el arrendatario, que debe pagar por esos pozos, deseará extraer el petróleo más lentamente, de modo que pueda economizar en la cantidad de ellos. Incluso puede perforar muy pocos pozos, ya que al decidir cuánto vale uno nuevo no considerará la porción de los ingresos que reciba el arrendador como una regalía. Por lo tanto, la mayoría de los arrendamientos de petróleo y gas contienen una cláusula "de desarrollo" que obliga al arrendatario a perforar un número razonable de pozos, es decir, justificado por los costos. Una interrogante interesante que ha surgido en la interpretación de tales cláusulas consiste en saber si el arrendatario, al calcular los costos de un pozo nuevo, podrá incluir no sólo sus costos de perforación y otros costos directos sino también la menor recaudación de los pozos antiguos, ya que el pozo nuevo agotará el yacimiento de donde se alimentan los pozos viejos y el pozo nuevo. La respuesta económica, para la que hay cierto apoyo judicial, es afirmativa, porque el agotamiento de los pozos existentes es un costo de oportunidad genuino del pozo nuevo.[57]

Hasta ahora hemos venido considerando la división vertical o temporal de un derecho de propiedad. Hay también una división horizontal. El ejemplo extremo es el derecho comunal, como en el pastizal que comparten va-

Rev., 289 (1983). El derecho de los fideicomisos se entiende mejor como un cuerpo especializado del derecho contractual al que se aplica en gran medida el análisis económico de los contratos (véase el capítulo siguiente). Por lo que toca a la economía del derecho de los fideicomisos, véanse, por ejemplo, Benjamin Ho y David Huffman, "Trust and the Law", en *Research Handbook on Behavioral Law and Economics* (Kathryn Zeiler y Joshua C. Teitelbaum [comps.], próxima publicación); Robert H. Sitkoff, "The Economic Structure of Fiduciary Law", 91 *Boston U. L. Rev.*, 1039 (2011) y "An Agency Costs Theory of Trust Law", 89 *Cornell L. Rev.*, 621 (2004); Frank H. Easterbrook y Daniel R. Fischel, "Contract and Fiduciary Duty", 36 *J. Law & Econ.*, 425 (1993); John H. Langbein, "The Contractarian Basis of the Law of Trusts", 105 *Yale L. J.,* 625 (1995). El derecho de los fideicomisos se discute en varios lugares de este libro; véanse las referencias del índice analítico.

[57] Véase un análisis económico de estos problemas en Stephen F. Williams, "Implied Covenants in Oil and Gas Leases: Some General Principles", 29 *U. Kan. L. Rev.,* 153 (1981).

rios agricultores. Los derechos comunales difieren sólo en grado de los no derechos, de modo que son ineficientes a menos que los costos del cumplimiento forzoso de los derechos individuales sean desproporcionados a los beneficios.[58] Por extraño que parezca, los derechos comunales son creados con frecuencia por individuos, aunque ello ocurre en circunstancias en las que se minimiza el problema de la ineficiencia. Por ejemplo, es posible que A deje un terreno a B y C, sus hijos, en una propiedad conjunta indivisa (un arrendamiento en común o un arrendamiento conjunto). B y C se encuentran formalmente en una situación muy similar a la de los habitantes de una sociedad que no reconoce los derechos de propiedad. Si B gasta dinero para reparar algunas estructuras de la propiedad, C participará igualmente en el valor de las reparaciones, y viceversa. Aunque sólo hay dos partes, se crea el conocido problema del monopolio bilateral. Pero éste está mitigado por la relación familiar; esperamos más cooperación entre personas unidas por lazos de afecto (en el capítulo v volveremos a ocuparnos de este punto).

Además, el derecho requiere que se acredite al inquilino conjunto el valor de cualquier mejora que haga en la propiedad hasta por la cantidad en que la mejora aumente su valor. Por otra parte, un inquilino conjunto que repara la propiedad tiene derecho a transferir la misma porción de los gastos de reparación (si los costos son razonables) al otro inquilino (o inquilinos) como parte de la propiedad de éste. Adviértase el contraste entre las mejoras y las reparaciones: el inquilino conjunto que mejora una propiedad es recompensado con los beneficios de la mejora, mientras que el inquilino conjunto que repara la propiedad sólo puede transferir una porción de los costos de las reparaciones al otro inquilino, ya que la reparación mantiene el valor de la propiedad más que aumentarla.

El derecho permite que un inquilino conjunto obtenga una partición de la propiedad en parcelas separadas de propiedad individual, lo que elimina las ineficiencias asociadas a los derechos comunales y el monopolio bilateral, mientras que a la vez protege las valuaciones subjetivas, lo que se perdería si la única forma de salir de un arrendamiento conjunto fuera una venta de la propiedad y una división de los fondos de la venta. Pero si hay economías de escala (dramáticamente ilustradas por el caso en que la propiedad conjunta es una pintura o una estatua), de modo que la propiedad valga menos dividida que entera, el tribunal rechaza la partición y ordena, en cambio, que se venda la propiedad.[59]

Supongamos que los propietarios adyacentes de casas en hilera, que comparten una pared medianera, no pueden ponerse de acuerdo sobre cómo

[58] Tenemos un buen ejemplo en el estacionamiento de un supermercado: no conviene (¡fuera de Manhattan!) cobrar a cada cliente por su uso de un espacio, aunque ello permitiría que el supermercado tuviera un estacionamiento ligeramente más pequeño.

[59] Véase Thomas J. Miceli y C. F. Sirmans, "Partition of Real Estate; or, Breaking Up Is (Not) Hard to Do", 29 *J. Legal Stud.*, 783 (2000). Véase más sobre la distribución de los fiduciarios entre los inquilinos a perpetuidad y los propietarios del resto en el capítulo xvi, problema 2.

se dividirán el costo de remplazar la pared, que está en peligro inminente de caerse. Uno de los propietarios la remplaza a su costa y luego demanda al otro por la mitad del costo. Hay una buena cantidad de autoridad judicial a favor de que se admita el juicio, como una solución para el problema del monopolio bilateral. Véase después IV.15.

El ingenio del derecho no es ilimitado, y terminaremos esta sección con un ejemplo deslucido de un caso de propiedad dividida acerca del cual el derecho no puede hacer nada: las rentas de automóviles. Como lo saben todos los que hayan rentado un automóvil, los individuos no tratan a los automóviles rentados con tanto cuidado como a los de su propiedad: son más rudos con ellos, lo que refleja el horizonte de tiempo acortado de su uso.[60] Pero en virtud de que la compañía arrendadora no puede supervisar o monitorear ese uso o determinar de manera factible el desgaste por el uso causado por un arrendatario en particular, no hay manera de que pueda inducir al arrendatario a tener el suficiente cuidado con el automóvil. Tenemos así un caso en el que los costos de transacción son elevados a pesar de que son pocas las partes y no hay problema de monopolio bilateral. La razón es que el costo del cumplimiento forzoso de la transacción convenida es prohibitivo porque los costos de información pertinentes son desproporcionados en relación con la inversión.

III.11. Problemas en la transferencia de los derechos de propiedad

La propiedad dividida dificulta la transferencia en la práctica, aunque no haya una limitación formal. Si 50 personas diferentes son inquilinos conjuntos de una propiedad, la venta de la propiedad requerirá que todos acepten el precio y la división del dinero entre ellos; habrá problemas de resistidores. Las elaboradas redes de parentesco de las sociedades primitivas impiden el surgimiento de los derechos de propiedad en ellas.[61] La eficiencia requiere que los derechos de propiedad sean transferibles; si muchas personas tienen un derecho sobre cada propiedad, las transferencias serán difíciles de arreglar.

La historia del derecho de la tierra en Inglaterra es una historia de esfuerzos para lograr que la tierra sea más fácilmente transferible. La "regla en el caso de Shelley" establecía que, si un otorgante concedía a A una posesión de por vida y el resto a los herederos de A, entonces A tenía un título simple (es decir, una propiedad absoluta); los herederos quedaban fuera. Si se reconocía el resto a los herederos de A, sería muy difícil transferir la propiedad, porque sus herederos no serían determinados hasta su muerte. La "doctrina del título más valioso" establecía que, si un otorgante concedía la propiedad a A

[60] Esto es lo contrario del problema de las mejoras del inquilino.
[61] Richard A. Posner, *The Economics of Justice,* 180-181 (1981).

de por vida, con el resto para los herederos del otorgante, éste —no sus herederos— era el dueño del resto. Así, podría venderlo, algo que sus presuntos herederos no podrían hacer a causa de la incertidumbre de su interés, porque el otorgante podría tener más hijos, quienes serían también sus herederos.

La objeción económica a estas doctrinas (además de su inmensa complejidad, no sugerida en estas descripciones) es que implica que el otorgante no puede recibir como compensación por los costos de una posibilidad de transferencia reducida cualesquiera beneficios que obtenga de la división de la propiedad en la forma que las doctrinas impiden. ¿Acaso los individuos no conocen sus intereses mejor que los tribunales? Pero estos otorgamientos son transacciones de una sola vez en la vida para el otorgante; y es posible que no tenga buena información acerca del problema que crean. (En el capítulo XIX volveremos a ocuparnos de este problema.) Además, los individuos que crean intereses excesivamente complejos abruman a los tribunales al igual que a sí mismos y a sus beneficiarios, de modo que hay externalidades que podrían justificar la intervención pública. Este punto explica la presunción del derecho común de que una transmisión de tierra a un ferrocarril u otra compañía de derechos de vía es la transmisión de un derecho de vía, es decir, un permiso de ingreso o servidumbre terminable cuando acabe el uso del adquirente, antes que la transmisión de un título simple. Los costos de transacción se minimizan con la propiedad indivisa de un terreno, y la propiedad indivisa se ve facilitada a su vez por la reunificación automática de la tierra dividida en cuanto cesa la razón de la división. Si el ferrocarril es propietario de una multitud de franjas de tierra estrechas que por ahora sólo puede usar el dueño de la tierra circundante o adyacente, antes de que la tierra pueda destinarse a su mejor uso deberá haber una negociación costosa y consumidora de tiempo o, en su defecto, la extinción gradual del interés del ferrocarril mediante la operación de la doctrina de la posesión adversa (que se discute más adelante). Es más limpia la eliminación del interés del ferrocarril cuando abandona el servicio ferroviario, aunque en algunos casos esto podría inducir al ferrocarril a continuar el servicio por más tiempo que el que es eficiente, con la esperanza de vender sus permisos de ingreso o servidumbres antes que abandonarlas.

Los derechos de agua son un ejemplo dramático de las externalidades que puede crear la transferencia de un derecho de propiedad.[62] Como se

[62] Véanse Charles J. Meyers y Richard A. Posner, *Market Transfers of Water Rights: Toward an Improved Market in Water Resources* (informe de la Comisión Nacional del Agua, 1º de julio de 1971, publicado por el Servicio Nacional de Información Técnica); Stephen F. Williams, "The Requirement of Beneficial Use as a Cause of Waste in Water Resource Development", 23 *Natural Resources J.,* 7 (1983); Ronald N. Johnson, Micha Gisser y Michael Werner, "The Definition of a Surface Water Right and Transferability", 24 *J. Law & Econ.,* 273 (1981); J. Mark Ramseyer, "Water Law in Imperial Japan: Public Goods, Private Claims, and Legal Convergence", 18 *J. Legal Stud.,* 51 (1989). Véase también después XXVII.8.

mencionó antes, en los estados occidentales de la Unión Americana se obtiene un derecho de propiedad desviando el agua de una corriente natural y usándola para riego u otros propósitos. Con el tiempo, una corriente será apropiada en su totalidad en el sentido de que el volumen total de la corriente es propiedad de varios usuarios en cantidades variables. Por ejemplo, A podría tener un derecho a tomar 10 pies cúbicos (0.283 metros cúbicos) por segundo, durante los meses de julio a diciembre, de un pozo en una localidad especificada; B podría tener un derecho a tomar 8 pies cúbicos (0.2264 metros cúbicos) por segundo en otra ubicación durante un periodo especificado, etc. Además, los derechos del agua se denominan por la fecha de adquisición (la fecha de la primera desviación, o de la apropiación), y en épocas de sequía, se raciona el abasto disponible de acuerdo con la prioridad de la apropiación.

Si A desea vender su derecho a X, y X planea usar el agua en el mismo lugar y en la misma forma que A, la transferencia no tendrá ningún impacto sobre los derechos de agua de los otros usuarios de la corriente. Pero supongamos que A y todos los demás usuarios presentes son agricultores que usan el agua que desvían para el riego, mientras que X, el comprador potencial del derecho de A, es un municipio. Entonces la transferencia afectará a los tenedores de derechos más abajo del punto de desviación de la corriente. En promedio, cerca de la mitad del agua que un agricultor desvía para riego se filtra de regreso a la corriente, y este flujo de retorno puede ser apropiado y se apropia por otros agricultores. Un municipio puede consumir un porcentaje mucho mayor del agua que desvía, y lo que no consume podría regresar a la corriente en un punto diferente, o podría fluir hacia una corriente diferente en su totalidad, si el municipio está ubicado en un sistema hidrológico diferente de aquel en el que se encuentra el agricultor cuyo derecho de agua ha comprado.

Si se omitieran los efectos sobre el flujo de retorno, muchas transferencias de agua reducirían el valor total. Supongamos que el derecho de agua de A vale 100 dólares para él y 125 para X, el municipio; pero mientras que A regresa la mitad del agua que desvía a la corriente, donde es usada por B, X retornará sólo un cuarto del agua que obtiene de A, y en un punto mucho más debajo de B, donde será apropiado por D. Y supongamos que B no vendería su derecho al flujo de retorno de A por menos de 50 dólares, mientras que D vendería su derecho al flujo de retorno del municipio por 10 dólares. El hecho de permitir que A venda su derecho de agua a X porque vale más para X que para A sería ineficiente, porque el valor total del agua sería menor en sus nuevos usos (el de X y el de D) —135 dólares— que en sus usos actuales (el de A y el de B), que es de 150 dólares.

La ley enfrenta este problema exigiendo que las partes demuestren que la transferencia no perjudicará a otros usuarios. En la práctica, esto significa que A y X (en nuestro ejemplo), a fin de completar su transacción, tendrían

que compensar a B por la pérdida del flujo de retorno de A; no lo harían; y la transacción fracasaría, como debe hacerlo bajo nuestros supuestos. Pero hay una deficiencia en esta solución. Todo nuevo flujo de retorno que genere el comprador no será de su propiedad; será un beneficio externo que las partes de la transacción no consideren. Sea que los valores para A, X y B sigan siendo 100, 125 y 50 dólares, respectivamente, pero sea ahora que el valor del flujo de retorno de X para D sea 60 dólares. Si se vende el agua, su valor (185 dólares) superará ahora a su valor en sus usos actuales (150 dólares), pero la ley requeriría que X pagara un mínimo de 150 dólares (100 dólares para A —el precio de reserva de A— más 50 dólares para B) por agua que vale sólo 125 dólares para él. X no será compensado por los 60 dólares de valor nuevo que cree su uso, de modo que se negará a completar la venta a menos que pueda inducir a D a pagar la diferencia existente entre 125 dólares y lo que les debe a A y a B. Para hacer esto, debe convencer a D de que no podrá realizar la compra sin tal contribución, ya que D sabe que podrá apropiarse del flujo de retorno de X, si se materializa, sin tener que pagar nada. Tenemos de nuevo un problema de monopolio bilateral. Una solución más eficiente, especialmente cuando puede beneficiarse más de un usuario por el flujo de retorno de nueva creación, consistiría en considerar a quien recibe la transferencia (X) como propietario de cualquier flujo de retorno nuevo que se cree por la transferencia.

Los problemas de la transferencia de derechos de propiedad forman parte de un problema más amplio: el de decidir quién es dueño de cuál propiedad. Un sistema de títulos registrados es una gran ayuda; y uno de los problemas de la transferencia de derechos de aguas es la ausencia de tal sistema adecuado. No sólo no se pueden adquirir derechos de aguas sino mediante el uso efectivo de la cantidad de agua reclamada, sino que el desuso durante cierto periodo de años conducirá a una pérdida de los derechos y su entrega al nuevo usuario. Una "escritura" de derechos de agua, que establezca que A tiene el derecho de tomar un número específico de pies cúbicos por segundo durante un periodo específico en un punto específico, es sólo una prueba del derecho legal de A y de lo que, en consecuencia, está facultado para vender. Y se requiere una investigación en el sitio para verificar que A, en efecto, posee (es decir, usa) lo que el registro de papel dice que es de su propiedad. Y a fin de determinar la significación de la prioridad de A en épocas de sequía, deben investigarse también los usos efectivos de otros apropiadores. Mejoraría la eficiencia si se instituyera un sistema de títulos de papel de aguas análogos a los sistemas usados para registrar los títulos de tierras.

Pero un sistema de registro no es una panacea, como lo demuestra la doctrina de la prescripción adquisitiva. Si durante un periodo dado de años (que es diferente en diversos estados de la Unión Americana, pero siete es un número común) mantiene usted la propiedad en perjuicio del propietario real (es decir, no como un inquilino, agente, etc., sino reclamándola como

suya), y él no demanda para ejercer su derecho, la propiedad será de usted. Holmes sugirió hace mucho tiempo una explicación económica para la prescripción adquisitiva. A lo largo del tiempo, una persona se vuelve apegada a la propiedad que considera como suya,[63] y la privación de la propiedad sería destructiva. Durante el mismo periodo de tiempo, una persona pierde su apego a la propiedad que ya no considera como suya, y por ello el restablecimiento de la propiedad causaría sólo un placer moderado.[64]

Esta observación se aplica en general a las prescripciones liberatorias (la adquisitiva es una transferencia efectuada por la prescripción liberatoria), aunque la prescripción adquisitiva tiene también la función más mundana de corregir los títulos. La mayoría de las prescripciones adquisitivas son resultado de errores causados por la incertidumbre acerca de las líneas divisorias. Las prescripciones liberatorias tienen también una función procedimental que analizamos en el capítulo XXII. Tales prescripciones disminuyen los costos de los errores causados por el uso de pruebas obsoletas para decidir una disputa.

Debiéramos considerar el impacto de la prescripción adquisitiva (o usucapión, el término más general para denotar la adquisición de derechos por el paso del tiempo) no sólo sobre el tenedor del título original y el posesionario, sino también sobre los compradores y los acreedores hipotecarios de la propiedad, quienes desean saber qué es exactamente lo que están comprando o tomando como garantía por el dinero que están prestando.[65] Por una parte, disminuyen sus costos de búsqueda porque pueden desentenderse de los gravámenes del título revelados sólo en los registros anteriores al periodo de la prescripción. Por otra parte, los registros ya no constituyen un índice completo de la propiedad. A fin de determinar el efecto neto de la prescripción adquisitiva sobre terceros, deberán ponderarse los costos de la inspección (o del aseguramiento del título) contra el ahorro de los costos de la búsqueda en registros muy antiguos.

[63] Éste puede no ser solamente un fenómeno psicológico. Creer que somos propietarios de alguna parte de una propiedad puede alentarnos a hacer inversiones complementarias, tales como conseguir un empleo en las inmediaciones, lo que podría tener que descartarse si perdemos la propiedad. Este punto se relaciona con la razón por la que es bastante probable que un individuo valúe su casa a un precio superior a su precio de mercado.

[64] Oliver Wendell Holmes, "The Path of the Law", 10 *Harv. L. Rev.,* 457, 477 (1897). Esta explicación implica que el posesionario se crea el propietario real. Hay un fuerte apoyo para este requerimiento en el derecho de casos particulares; R. H. Helmholz, "Adverse Possession and Subjective Intent", 61 *Wash. U. L. Q.,* 331 (1983). La teoría de Holmes se replantea como una teoría acerca de la protección del interés de confianza del posesionario en Thomas J. Miceli y C. F. Sirmans, "An Economic Theory of Adverse Possession", 15 *Int'l Rev. Law & Econ.,* 161 (1995). Como sabemos por nuestra discusión del dominio eminente, y como veremos de nuevo en el siguiente capítulo, la protección legal de los intereses de confianza puede conducir a una confianza excesiva. El peligro se limita en el contexto de la prescripción adquisitiva por el riesgo de que el propietario original afirme sus derechos dentro del periodo estatutario, lo que destruye la confianza del posesionario.

[65] Robert C. Ellickson, "Adverse Possession and Perpetuities Law: Two Dents in the Libertarian Model of Property Rights", 64 *Wash. U. L. Q.,* 723, 730-731 (1986).

Entre menor sea el periodo de la prescripción, habrá menos incertidumbre en lo tocante a los títulos en las transferencias posteriores de la propiedad, pero mayor será el peligro de que el dueño pierda la propiedad por accidente,[66] al igual que mientras menor sea el periodo de limitaciones, mayor será el riesgo de perder el derecho por falta de diligencia; mientras que entre mayor sea el periodo de prescripción, más probable será que se determine de manera errónea quién es el verdadero dueño (al igual que un plazo muy largo de limitaciones fomenta el retraso que podría impedir una determinación precisa de los derechos) y también será más dramático el "efecto de Holmes" con el que iniciamos esta discusión.[67]

No queda claro si es el plazo más corto o el más largo el que minimiza los errores ponderados por su costo, pero existen pruebas de que entre mayores sean la densidad de población y los valores de la propiedad, el periodo de prescripción será menor; y se ha dado la explicación de que los errores son más frecuentes cuando hay una mayor densidad y más costosos cuando hay un valor mayor en juego; y, como se sugirió anteriormente, un plazo menor de prescripción minimiza la cantidad de errores.[68]

El requisito de la adversidad en el derecho de la prescripción adquisitiva merece énfasis. Debe ser evidente que alguien está ocupando la propiedad del dueño nominal en virtud de una reclamación de derecho no derivado de ese propietario: el tipo de reclamación que tendría un inquilino.

La discusión tiene implicaciones para la elección entre los tipos de sistemas de registro.[69] Bajo el sistema de registro empleado en la mayoría de los estados, en el caso de una interrupción en la cadena de títulos a resultas de la cual el ocupante actual no es el verdadero dueño (quizás el registrador de escrituras no registró una escritura que algún dueño anterior había extendido), la tierra se devuelve al verdadero dueño. El ocupante actual se protege contra tal posibilidad comprando un seguro de título, y así recibe una compensación monetaria por su pérdida. Bajo el sistema de Torrens vigente en algunos estados de la Unión Americana, el ocupante actual prevalece sobre el dueño verdadero, a quien se compensa por su pérdida con un fondo público. El enfoque de la mayoría de los estados es más eficiente por cuanto recurre al seguro privado antes que al social para compensar al reclamante desposeído, pero el sistema de Torrens es más eficiente pues permite que el ocupante actual, quien probablemente valúa más la propiedad (el criterio de Holmes), la conserve sin incurrir en el costo de transacción de su compra al dueño verdadero.

[66] El riesgo es particularmente grave para los propietarios de los bienes personales perdidos, lo que podría tomar mucho tiempo para su localización.

[67] Véase un análisis empírico en Jeffrey M. Netter, Philip L. Hersch y William D. Manson, "An Economic Analysis of Adverse Possession Statutes", 6 *Int'l Rev. Law & Econ.,* 217 (1986).

[68] Véase Ben Depoorter, "Adverse Possession", en *Property Law and Economics*, 183 (Boudewijn Bouckaert [comp.], 2010).

[69] Véase Thomas J. Miceli y C. F. Sirmans, "The Economics of Land Transfer and Title Insurance", 10 *J. Real Estate Finance & Econ.,* 81 (1995).

Más adelante la doctrina de los compradores de buena fe ilustra la transferencia forzada. A le entrega su abrigo a su agente B para que lo empeñe, y B, entendiendo mal las instrucciones, lo vende a C. Siempre que C no sepa o no tenga razón para saber que B no estaba autorizado para venderle el abrigo, C adquirirá un título bueno sobre el abrigo. Éste es un caso simple en el que A puede evitar el error a un costo menor que C (aunque B es quien puede evitar el error al costo más bajo en comparación con los demás y A podría obtener de él una indemnización, de tal modo que la pérdida se transferiría de C hacia A y de éste hacia B). Pero supongamos ahora que B no era un agente de A, sino que se robó el abrigo y lo vendió a C sin dar a C ninguna razón para sospechar que estaba comprando mercancía robada. C no adquirirá un título bueno; un ladrón no puede transferir un título bueno a su comprador. Aunque A podría prevenir la transferencia errónea a un costo menor que C al tomar precauciones mayores contra el robo, si se permitiera que C obtuviera un título bueno se estimularía el robo. Los ladrones obtendrían precios mayores de sus encubridores, porque los encubridores podrían obtenerlos (siempre que tomaran medidas para que el comprador no sospechara respecto a su origen) en el mercado de reventa; la gente pagará más por un título seguro que por uno confuso. No queremos un mercado eficiente de bienes robados.

Sin embargo, puede argüirse que debiera existir una excepción para los bienes duraderos, en particular las obras de arte.[70] Se robaron muchas obras de arte durante la segunda Guerra Mundial, que terminó hace más de 70 años; si el dueño original no ha hecho nada para tratar de recuperar la obra durante todo ese tiempo, ¿no debiera excluírsele? ¿No existe acaso el peligro de que, si el dueño de una obra de arte no puede estar nunca seguro de que un dueño anterior no saldrá de la nada para recuperar la obra, la gente se resistirá a exhibir sus obras de arte por temor a alertar a un dueño anterior? Pero la otra cara de esta moneda es que los dueños originales tomarán precauciones adicionales para impedir el robo de sus obras de arte —y estas precauciones pueden incluir la exhibición discreta— si saben que no podrán recuperar el arte robado y se vende a un comprador de buena fe. El costo de las precauciones adicionales para el dueño, si no puede recuperar la obra de arte que le robaron, tiene que balancearse contra el costo de los esfuerzos adicionales del comprador por impedir el descubrimiento de la obra de arte y los costos adicionales de la búsqueda de los dueños originales para descubrir su arte robado, si éstos pueden recuperarlo. Si los costos del ocultamiento del comprador y de la búsqueda del dueño, bajo el sistema de que el dueño original gana, no superan grandemente a los costos de la precaución del dueño bajo el sistema de que el comprador de buena fe gana, entonces la inconveniencia de

[70] Esta cuestión se discute en William M. Landes y Richard A. Posner, "The Economics of Legal Disputes over the Ownership of Works of Art and Other Collectibles", en *Essays in the Economics of the Arts,* 177 (Victor A. Ginsburgh y Pierre-Michel Menger [comps.], 1996).

que los bienes robados se vuelvan más fácilmente vendibles en el mercado debiera inclinar la balanza contra la aceptación del principio de que quien compra de buena fe a un ladrón adquiere el título.

El problema implicado en el caso del robo de obras de arte se asemeja al de la duración óptima del periodo de la prescripción. Entre mayor sea el periodo, más seguro será el título del dueño original: tendrá más tiempo para descubrir que alguien mantiene una prescripción adquisitiva en su contra. Así, en igualdad de otras circunstancias, entre mayores sean los costos del monitoreo de nuestros derechos, más largo debiera ser el periodo de la prescripción.[71] Aquí se trata de abordar el problema desde el punto de vista del rival, el propietario original, antes que del punto de vista del propietario actual. Un periodo de prescripción mayor ayuda al primero, así como un periodo de limitaciones más largo ayuda a la persona que afirma que sus derechos han sido vulnerados.

Además de la prescripción y de las limitaciones a los derechos de propiedad intelectual (las patentes y la mayoría de los derechos de autor tienen una duración limitada), en general, los derechos de propiedad no tienen una duración limitada. Sin embargo, bajo la norma del derecho común contra las perpetuidades (la cual, no obstante, ha sido modificada o suprimida en la mayoría de los estados) no podría crearse un derecho de propiedad que surta efectos por más de 21 años después de la muerte de una persona que vivía cuando supuestamente se otorgó el derecho. La norma refleja los problemas de la "mano muerta", principalmente relacionados con los testamentos y los fideicomisos, los cuales discutimos en el capítulo XIX. Pero en ocasiones surge en casos de propiedad convencionales, como en American Land Holdings of Indiana, LLC *vs*. Jones.[72] Una escritura de 1903 traspasó a un predecesor de la empresa de carbón Peabody el derecho a explotar el carbón existente debajo de la tierra del demandado. Ese derecho no se hubiera perdido, incluso si la explotación del carbón no hubiera comenzado hasta 2000. La razón es que el derecho a la explotación es una servidumbre accesoria: una servidumbre o permiso de ingreso (lo cual puede otorgarse de forma expresa, como en la escritura involucrada en el caso, o por implicación, como cuando un terrateniente vende una parcela totalmente rodeada por su tierra y se considera que el comprador tiene una servidumbre implícita de entrada y salida a través de la propiedad del vendedor) que se ejecuta con la tierra porque es necesaria para la explotación óptima de otro derecho de propiedad y, por lo tanto, satisface la condición de una servidumbre de ese tipo, que es la de aumentar el valor de la propiedad de la cual es accesoria. El derecho a la explotación del carbón es

[71] Véase Matthew Baker, Thomas Miceli, C. F. Sirmans y Geoffrey Turnbull, "Property Rights by Squatting: Land Ownership Risk and Adverse Possession Statutes", 77 *Land Econ.,* 360 (2001); Omri Ben-Shahar, "The Erosion of Rights by Past Breach", 1 *Am. Law & Econ. Rev.,* 190, 223-225 (1999).

[72] 604 F.3d, 451 (7th Cir., 2010).

accesorio a la titularidad de un depósito de carbón, ya que sin ese derecho el valor del carbón disminuiría para su dueño… aunque no a cero, porque el propietario de la superficie pagaría algo, tal vez una gran cantidad, para poder extraer el carbón que está debajo de su tierra.

Condicionar el ejercicio de un derecho accesorio a la Regla contra las Perpetuidades alentaría la explotación prematura, un tema que se discute a detalle en la siguiente sección. Supongamos que luego de vender los derechos sobre el carbón al predecesor de Peabody, en 1903, el precio del carbón cayó o que el costo de su extracción se disparó y en consecuencia la extracción del carbón era poco rentable, pero que las condiciones mejoraron gradualmente y, en 1923, la empresa de carbón determinó que la extracción del carbón sería rentable a partir de 1925. Si para conservar su derecho a explotar la empresa de carbón tuviera que empezar la minería a los 21 años de la adquisición del derecho, tendría un incentivo para comenzar la explotación antes de tiempo, con el fin de preservar su derecho. Eso sería un desperdicio. En efecto, la compañía estaría explotando para adquirir el derecho a extraer en el futuro, en lugar de explotar porque quería extraer y vender el carbón ahora.

III.12. La posesión

Podemos reunir varios de los puntos discutidos en este capítulo concentrándonos más sistemáticamente en el concepto fundamental de la *posesión* como una fuente de los derechos de propiedad, ya tocado en varios lugares en este capítulo; por ejemplo, con referencia a los animales salvajes, los tesoros y la prescripción.[73]

Podemos imaginar dos sistemas polares de derechos de propiedad exactamente opuestos: la propiedad de acuerdo sólo con un sistema de títulos de papel, y la propiedad sólo por la posesión física. Ambos sistemas implicarían graves ineficiencias. Un sistema universal de títulos de papel supone que todo es ya propiedad de alguien[74] y sólo permite las transferencias por la transmisión formal (por ejemplo, la entrega de una escritura), de modo que es inútil buscar la adquisición de propiedades que no tengan dueño, ya sea porque nunca lo tuvieron o porque han sido abandonadas. Tal sistema dejaría también sin definición el estatus de los no propietarios que, sin embargo, tienen el uso exclusivo de la propiedad, como ocurre con los inquilinos. Y sería inútil examinar los errores inevitables que causa un sistema de derechos de papel. El otro régimen extremo, en el que los derechos al uso exclusivo de la propiedad se hacen depender del control físico, implica fuertes inversiones en el

[73] Por lo que toca a la economía de la posesión en general, véanse Richard A. Epstein, "Possession", en *The New Palgrave Dictionary of Economics and the Law*, vol. 3, p. 62 (Peter Newman [comp.], 1998); Dean Lueck, "First Possession", en *The New Palgrave…*, vol. 2, p. 132.

[74] Más adelante se discute una excepción: la adquisición del título por un donativo.

mantenimiento de tal control. Como sabemos, tampoco hace ninguna provisión para los derechos del uso futuro por oposición al uso presente.

Por lo tanto, es probable que un régimen jurídico de los derechos de propiedad eficiente sea un sistema mixto, que combine los derechos de papel con los derechos posesorios. Supongamos que se descubriera un continente nuevo, deshabitado. Si se diera al descubridor un título sobre todo el continente antes de que hubiese tomado posesión de él, en el sentido de ocuparlo todo o por lo menos en su mayor parte, se incitaría una inversión excesiva en exploración, ya que el explorador que descubriera el continente justo un día antes que sus rivales obtendría el valor entero del continente. Un caso más extremo aún, que fue común en el periodo inicial de la exploración europea de otros continentes, fue el esfuerzo de los monarcas (incluido el papa) para crear, por donativos, derechos de propiedad en tierras no descubiertas aún.

La alternativa de basar en la ocupación física la propiedad de tierras sin dueño anterior disminuye la recompensa neta de ser el primero y, así, alivia el problema de la inversión excesiva, obligando al dueño potencial a incurrir en costos de ocupación. También tiende a asignar recursos a las personas que puedan usarlas más productivamente, porque ellas son quienes tienen mayores probabilidades de estar dispuestas a incurrir en los costos implicados en la posesión. Un descubridor que pudiera obtener un título para todo el continente sólo por una declaración o un registro pronto vendería la mayor parte o el total de la tierra, porque no sería el desarrollador más eficiente de toda ella. Es más eficiente otorgar el derecho de propiedad en primer lugar a la gente que vaya a poseer efectivamente la tierra.

¿Pero qué significa exactamente "posesión" en este contexto? Supongamos que el descubridor y primer poseedor de un terreno que no tenía dueño previamente, que nadie reclamaba ni ocupaba, se ausenta continuamente. Si alguien ocupa ahora la tierra, ¿es él el poseedor? En tal caso, los dueños harían gastos dispendiosos en el cercado y el patrullaje de su tierra. Una cosa es condicionar la adquisición del título a la posesión de la propiedad recién encontrada, pero, una vez que se adquiere el título por esta vía, debiera bastar para el mantenimiento de éste su registro en una oficina pública de escrituras a fin de alejar a los allanadores accidentales. Ése es un método de aviso más barato que la elaborada colocación de señales y de cercas, ya no digamos la clase de uso presente, generalizado, que podría requerirse razonablemente para obtener el título de un continente recién descubierto.

Consideremos el caso siguiente, a primera vista sin ninguna relación: el demandante le encarga una caja fuerte al demandado para que la venda por su cuenta. El demandado encuentra algo de dinero, evidentemente propiedad del demandante, en un rincón de la caja fuerte. El demandante exige el regreso del dinero. ¿Debiera recuperarlo? El encuentro de una propiedad perdida es un servicio valioso y debiera alentársele. Pero, al igual que ocurre con el descubrimiento de continentes nuevos, si se entrega al descubridor

todo el valor de la propiedad podría generarse una inversión excesiva en exploración. Y también podría suceder que los dueños invirtieran en exceso en la salvaguardia de su propiedad. Es mejor entregar al descubridor, por el hallazgo, una recompensa que es el dominio del derecho de la restitución (véanse después IV.9 y VI.9), antes que la propiedad de la cosa encontrada.

Supongamos que una persona deja su cartera, que contiene dinero, en el mostrador de un supermercado. Otro cliente recoge la cartera. El dueño nunca la reclama. ¿Debiera el cliente tener derecho a conservar la posesión de la cartera y el dinero, o es el supermercado (el *locus in quo,* como se dice en los juicios) el que tiene ese derecho? En otras palabras, ¿la "posesión" de quién determinará la propiedad? El argumento a favor del cliente es que, puesto que fue él quien la encontró, merece una recompensa; el supermercado no hizo nada. Pero si, sabiendo que podrá quedarse con la cartera a menos que el dueño la reclame, el cliente se sale de la tienda con ella, es menos probable que se devuelva al dueño que si la deja en la tienda para que un empleado la encuentre. Porque cuando el dueño descubra su pérdida preguntará en los lugares en los que ha estado, y la búsqueda lo llevará de regreso al supermercado.

El derecho común trata tales casos distinguiendo entre las cosas perdidas y las extraviadas, en donde "perdidas" significa que el dueño no se da cuenta de que el bien no se encuentra. Si no lo advierte, es improbable que busque la propiedad, de modo que la ley convierte al descubridor en el poseedor legal y no al dueño del lugar donde se encuentre, como sucede en el caso de los bienes extraviados. (El caso de la caja fuerte involucra una propiedad perdida, y el caso de la cartera una propiedad extraviada.) Pero hay una objeción económica a que se le permita al cliente-descubridor quedarse aun con los bienes perdidos, por oposición a los extraviados: su recompensa podría superar a su costo por mucho, y hemos visto que las recompensas excesivas por el hallazgo tienden a atraer recursos excesivos hacia las actividades que generen tales recompensas. Es cierto que el cliente-descubridor sólo obtiene esta recompensa *ex post*; es decir, sólo si el dueño no reclamó su propiedad. Y esto significa que la recompensa *esperada* por el descubridor podría haber sido pequeña, ya que la mayoría de la gente que pierde una propiedad valiosa hace un esfuerzo para recuperarla. Pero, dado que un empleado del supermercado habría encontrado probablemente la cartera poco tiempo después de que lo hiciera el cliente, el valor del hallazgo del cliente podría haber sido escaso… de hecho, negativo, porque el dueño tendrá más dificultad para reclamarla de un cliente que del supermercado, aunque el cliente deba dejar su nombre y domicilio en el supermercado.

Por lo tanto, supongamos que la regla es que el supermercado tiene una posesión legal pero el cliente-descubridor no sabe o no le importa lo que diga la ley y se marcha con la cartera, y luego la olvida en el siguiente supermercado que visita. Ahora la encuentra un empleado del supermercado y el

cliente vuelve al supermercado y la reclama. ¿Debiera prevalecer el cliente sobre el descubridor subsecuente, el supermercado? No, porque queremos disuadir al cliente-descubridor de quitar la propiedad al primer descubridor, el supermercado.

El caso de la caja fuerte plantea la cuestión de si debiera requerirse el control físico para el mantenimiento de la posesión legal, así como para su adquisición. En general, la respuesta de la economía es negativa. Tal requerimiento conduciría a gastos dispendiosos y también desalentaría la especialización. Imaginemos que se considerara a un inquilino como dueño de las instalaciones arrendadas porque el terrateniente, en virtud del contrato de arrendamiento, ha perdido su control físico (es decir, el terrateniente no puede entrar a las instalaciones durante el periodo del arrendamiento). Sería más razonable reconocer la posesión conjunta del terrateniente y el inquilino, y dividir el derecho de ejercer una acción legal para proteger entre ellos sus intereses posesorios, de acuerdo con la ventaja comparativa en circunstancias particulares. Estas circunstancias variarán. Comparemos los casos siguientes: *1)* el inquilino no ha tomado aún la posesión; *2)* el despojo realizado por un intruso ocurre tan tarde en el periodo del arrendamiento que el inquilino tiene escaso incentivo para demandar; *3)* la infracción es más dañina para el terrateniente que para el inquilino (por ejemplo, si el inquilino es despojado por un narcotraficante, quien procede a atemorizar a los otros inquilinos para que se marchen), y *4)* el inquilino carece simplemente de los recursos necesarios para litigar contra el infractor.

¿Cuándo debiera considerarse no perdida ni extraviada sino abandonada la propiedad, es decir, devuelta a la comunidad de recursos sin dueño y, por ende, disponible para su apropiación por alguien más? Desde un punto de vista económico, el caso más claro del abandono ocurre cuando el dueño "tira" deliberadamente la propiedad, devolviéndola por voluntad a la comunidad. Su acto significa que la propiedad no tiene valor para él, de modo que al considerar abandonada la propiedad y, por ende, disponible para su reapropiación por alguien más la ley estimula la reasignación de la propiedad a un uso de mayor valor sin cargar al sistema con costos de negociación. Porque dado que el dueño (anterior) la valúa en cero dólares o menos, *quienquiera* que se moleste en tomar la propiedad la valuará seguramente en más. Además, los costos de transacción pueden ser elevados, incluso cuando se esté tratando con parcelas de tierra. El dueño puede ser desconocido. Para cuando el dueño despierta y afirma sus derechos, es posible que las pruebas hayan desaparecido y que el nuevo posesionario —sin saberlo, un posesionario adversario— se haya basado en una creencia razonable de que él es el verdadero dueño. Pensando que la propiedad es suya, puede que haya hecho en ella una inversión que no valdrá nada si pierde la propiedad ante el dueño original, para quien, sin embargo, la propiedad puede no valer nada, como lo indicó al dejar "dormidos" sus derechos. Cuando hay una gran disparidad

en el valor que los únicos competidores por un bien le asignan, los costos de transacción tenderán probablemente a ser elevados, porque cada competidor se esforzará para obtener la mayor parte posible de ese valor.[75] Así, pues, la prescripción adquisitiva es un método para la corrección de los títulos de papel en contextos en los que los costos de transacción del mercado son elevados;[76] este método mejora al sistema de los derechos de propiedad en lugar de cuestionarlo.

Justo como una escritura de título inscrita en un registro público, la posesión es un procedimiento de notificar la existencia de un derecho a todo el mundo, siempre que sea "abierta y notoria", como dicen los fallos sobre la posesión adversaria.[77] Es el único procedimiento viable en las primeras etapas de la sociedad; la cerca es anterior al título de propiedad como un método de anuncio de un derecho de propiedad. El requisito de un ejercicio del poder físico para la obtención o el mantenimiento de un derecho posesorio compensa los costos de actos físicos particulares que comunican una reclamación con los beneficios de una comunicación clara. Entre más refinados sean los actos requeridos, más exenta de errores estará la comunicación (lo que es bueno porque la definición pública clara de los derechos de propiedad reduce los costos de transacción y tiende a optimizar la inversión); pero también se vuelve más costosa esta forma de la notificación. Los costos de los actos más refinados de la notificación por la posesión —actos de ocupación completa, continua y conspicua— superarán a menudo a los beneficios. Por eso un grado menor de la posesión activa bastará para mantener un derecho de propiedad de lo que sería requerido para adquirirlo.

Consideremos el colorido caso antiguo de Haslem *vs.* Lockwood.[78] El demandante había recogido en montones la defecación de caballos tirada en las calles que intentaba llevarse al día siguiente, lo más pronto que podría conseguir la transportación necesaria. El demandado se apoderó del montón y lo transportó para venderlo. El demandante pidió la devolución del fertilizante y ganó. Los dueños originales del fertilizante, que eran los dueños de los caballos que lo habían expulsado, lo habían abandonado; el demandante se lo había encontrado. Tomó posesión del fertilizante al juntarlo en montones, y los montones eran un aviso adecuado para los terceros, como el de-

[75] Supongamos que la tierra vale un millón de dólares para el posesionario adversario (quizá porque sabe que hay depósitos minerales en ella) y sólo 10 000 dólares para el dueño original (quien ignora la existencia de los depósitos). Entonces, a cualquier precio entre 10 000 y un millón de dólares, ambas partes estarán en mejor situación por una venta. Pero ambas partes estarán ávidas de obtener la mayor parte posible de la diferencia, y eso podría dificultar para ellas un acuerdo sobre el precio sin una negociación prolongada y costosa.

[76] Thomas W. Merrill, "Property Rules, Liability Rules, and Adverse Possession", 79 *Nw. U. L. Rev.,* 1122 (1985).

[77] Esta función de la posesión se destaca en Carol M. Rose, "Possession as the Origin of Property", 52 *U. Chi. L. Rev.,* 73 (1985).

[78] 37 *Conn.,* 500 (1871).

mandado, de que el fertilizante (ya no) estaba abandonado. Si se hubiese requerido que el demandante, a fin de proteger su derecho de propiedad, hiciese algo más que reunir el fertilizante en montones —que lo hubiese cercado, o lo hubiese vigilado continuamente, o hubiese hecho arreglos por adelantado para tener en el lugar un carro que se llevara el fertilizante en cuanto fuera amontonándose—, habría incrementado el costo de la transacción por la que el fertilizante sin valor para el dueño original se convirtió en un bien valioso, sin generar beneficios compensatorios.

Algunos estados tratan de determinar o restaurar la propiedad perdida o extraviada, a diferencia de la propiedad abandonada (es decir, "botada" deliberadamente), al establecer que, después de un periodo determinado de años, la propiedad desocupada, no reclamada (mueble o inmueble) pasa a manos del Estado (es decir, se convierte en la propiedad de éste). Un Estado también podría establecer que, después de un periodo mucho más corto en el que se desconoce al dueño de la propiedad, el Estado deberá asumir la custodia de la misma, tratar de encontrar al dueño y, si el propietario aparece para reclamarla antes de que haya transcurrido el tiempo suficiente para que ésta pase a manos del Estado, devolver la propiedad al dueño menos un cargo por el costo de custodiarla y por la búsqueda del dueño. El objetivo de estos procedimientos es evitar situaciones en las que, a falta de un propietario identificado, el valor de la propiedad no se maximiza.

III.13. Los efectos distributivos de las asignaciones del derecho de propiedad

El economista puede auxiliar al encargado de formular las políticas no sólo explicando los efectos de una política sobre la eficiencia del uso de los recursos, sino también rastreando sus efectos sobre la distribución del ingreso y la riqueza. Consideremos una propuesta de que se haga responsable al dueño de una fábrica por el daño que sus humos causan a los dueños de propiedades residenciales. En un análisis superficial, el único efecto sobre la riqueza consiste en mejorar la situación de los casatenientes y empeorar la situación del dueño de la fábrica, seguramente un hombre rico. Pero el asunto es más complicado. Si el monto del daño causado por el humo (y el costo de diversas medidas para reducirlo) aumenta a medida que incrementa la producción, la nueva responsabilidad aumentará el costo marginal de la fábrica. ¿Podrá la empresa compensarse aumentando el precio de su producto? No lo hará si antes estaba vendiendo a un precio igual a su costo marginal y tiene competidores que venden a un precio igual a su costo marginal, que es el mismo que el suyo salvo que no son responsables de daños causados por el humo (o quizá tengan maquinaria más nueva que no produce humo como un subproducto). Si esta empresa tratara de aumentar su precio, sus clientes se irían con sus

competidores. Pero si la curva del costo marginal de la empresa es creciente, lo que quiere decir que su costo marginal es menor entre menor sea su producción, entonces podrá mantenerse a flote reduciendo su producción, a pesar de que los costos del control de la contaminación han desplazado hacia arriba su curva del costo marginal. Aun así, las ganancias de la empresa disminuirán, al igual que el número de sus empleados, el monto de los abastos que compra y el monto de la renta que puede pagar por la tierra y otros recursos escasos.

Supongamos ahora que todas las empresas rivales son responsables del daño causado por el humo, de modo que todas ellas experimentan un aumento de sus costos marginales igual al de la primera empresa. Un aumento del precio es ahora viable para la empresa. Sus ventas no bajarán a cero. Podemos suponer que el producto es idéntico para todas las empresas, pero no es idéntico a otros productos, de modo que los consumidores pagan más antes que prescindir del producto. Pero sabemos por el capítulo I que habrá *alguna* sustitución y, por lo tanto, bajará la producción de la industria. Pero los consumidores comparten ahora la carga de la responsabilidad con los proveedores de insumos, porque algunos consumidores se cambian a otros productos cuando habrían preferido continuar comprando el producto de la industria a su precio anterior, mientras que otros continúan comprándolo y pagan un precio mayor.

III.14. Las tierras públicas

Gran parte de la tierra del oeste de los Estados Unidos es propiedad del gobierno federal, que la administra. El impulso original para la retención de extensas tierras públicas provino del movimiento conservacionista, temeroso de que el desarrollo privado condujera al agotamiento prematuro de los recursos naturales, principalmente la madera. No hay ninguna base económica para tal preocupación. Al decidir si se talará un árbol, el propietario privado de la tierra donde crece el árbol considerará no sólo el ingreso de la venta de la madera, el costo de la tala y el corte del árbol, sino también el costo de oportunidad de no esperar hasta que el árbol haya crecido a su máxima altura.

La administración gubernamental del recurso ha sido perversa. El gobierno limita el corte en cada área de tierra gubernamental al número de árboles que se han añadido por el nuevo crecimiento desde la última tala, a fin de impedir una reducción neta del número de árboles en el área. Los bosques muy viejos tienden a tener escaso crecimiento nuevo, porque los árboles están hacinados. Por lo tanto, se permite que pocos árboles sean talados en tales áreas… aunque una tala mayor, al permitir un crecimiento más rápido, podría maximizar el rendimiento del bosque a largo plazo. A la inversa, esta "regla del aumento" podría conducir a una tala excesiva de los bosques jóvenes.

La justificación principal de la extensa propiedad gubernamental de la tierra es estética: la preservación de áreas silvestres para disfrute de los campistas. La fundamentación económica es la dificultad de cobrar por el acceso, y es una justificación débil. Pocos caminos conducen a estas áreas, y las casetas de cobro podrían establecerse fácilmente en los puntos de entrada. De igual modo, los parques nacionales podrían ser privatizados; el argumento en favor de los parques públicos urbanos es más fuerte porque limitar el acceso es más costoso. Si por alguna razón se desea subsidiar el uso de tales instalaciones (aunque muy pocos de los usuarios son pobres), esto puede lograrse mediante un donativo de dinero al propietario privado; no hay ninguna razón económica convincente para la propiedad gubernamental.

¿Qué diremos de las personas que podrían desear visitar algún día los bosques nacionales y están dispuestos a pagar algo por esta opción, por improbable que sea que lleguen a aprovecharla? ¿Y qué diremos de las personas que creen que cualquier disminución de la cantidad de bosques podría tener efectos adversos a largo plazo sobre el clima y están dispuestas a apoyar esta creencia con dinero? ¿No debieran contar estas preferencias, estas preocupaciones? Por supuesto que sí. Si las personas están dispuestas a pagar impuestos para apoyar la existencia de más árboles que los que preservaría un mercado libre (en lugar de que estuvieran dispuestas simplemente a lograr que otras personas fueran involuntariamente gravadas para ese fin), porque los árboles generan beneficios externos al igual que apropiables, eso está muy bien. Pero debe distinguirse entre la escala y la propiedad: los bosques pueden ser subsidiados públicamente sin ser de propiedad pública.

Las encuestas de actitudes hacia los parques nacionales y otras tierras públicas recreativas revelan dramáticos efectos de dotación. Si se pregunta a la gente cuánto dinero demandaría por vender sus derechos al uso de tales tierras, da cifras mucho mayores que si se le preguntara cuánto ofrecería para comprar tales derechos. No debe pensarse que esa disparidad sea irracional. Es posible que refleje el carácter único del bien en cuestión, único en el sentido económico de la ausencia de sustitutos cercanos.[79] La ausencia de un sustituto cercano de un bien implica que éste no podría ser remplazado fácilmente si se perdiera, de modo que el propietario demanda un precio elevado por desprenderse de él. Pero si no es su propietario quizá no desee pagar un precio elevado por ese bien, pues quizá no sepa lo que se está perdiendo, ya que por hipótesis nada de lo que tiene se parece mucho a ese bien.[80]

[79] Véase Daniel S. Levy y David Friedman, "The Revenge of the Redwoods? Reconsidering Property Rights and the Economic Allocation of Natural Resources", 61 *U. Chi. L. Rev.,* 493 (1994).

[80] Más adelante examinaremos varios temas que se estudian a menudo en los cursos del derecho de la propiedad. Véanse XVII.8 (cumplimiento forzoso de los códigos de vivienda), XVIII.4 (impuestos prediales) y XIX.2-XIX.7 (herencias).

Lecturas recomendadas

Baxter, William F., *People or Penguins: The Case for Optimal Pollution* (1974).
Calabresi, Guido, y A. Douglas Melamed, "Property Rules, Liability Rules, and Inalienability: One View of the Cathedral", 85 *Harv. L. Rev.,* 1089 (1972).
Clawson, Marion, *The Economics of National Forest Management* (1976).
Coase, Ronald H., "The Problem of Social Cost", 3 *J. Law & Econ.,* 1 (1960).
Demsetz, Harold, "Toward a Theory of Property Rights", 57 *Am. Econ. Rev. Papers and Proceedings,* 347 (mayo, 1967).
———, "When Does the Rule of Liability Matter?", 1 *J. Legal Stud.,* 13 (1972).
Dnes, Antony, y Dean Lueck, "Asymmetric Information and the Law of Servitudes Governing Land", 38 *J. Legal Stud.*, 89 (2009).
Eckert, Ross D., *The Enclosure of Ocean Resources: Economics and the Law of the Sea* (1979).
Ellickson, Robert C., "Alternatives to Zoning: Covenants, Nuisance Rules, and Fines as Land Use Controls", 40 *U. Chi. L. Rev.,* 681 (1973).
Kaplow, Louis, y Steven Shavell, "Property Rules *versus* Liability Rules: An Economic Analysis", 109 *Harv. L. Rev*., 713 (1996).
Lueck, Dean, "The Rule of First Possession and the Design of the Law", 38 *J. Law & Econ*., 393 (1995).
Merrill, Thomas M., "The Economics of Public Use", 72 *Cornell L. Rev.,* 61 (1986).
———, "Trespass, Nuisance, and the Costs of Determining Property Rights", 14 *J. Legal Stud*., 13 (1985).
———, "Accession and Original Ownership", 1 *J. Legal Analysis,* 459 (2009).
Miceli, Thomas J., C. F. Sirmans y Geoffrey K. Turnbull, "The Property-Contract Boundary: An Economic Analysis of Leases", 3 *Am. Law and Econ. Rev.*, 165 (2001).
Smith, Henry E., "Property and Property Rules", 79 *N.Y.U. L. Rev.,* 1719 (2004).
Property Law and Economics (Boudewijn Bouckaert [comp.], 2010).

Problemas

1. Supongamos que puede saber si hay un depósito de petróleo bajo la tierra de alguien dirigiendo una onda de radio bajo la superficie de su tierra. No tiene que estar directamente sobre el depósito, es decir, dentro de sus líneas de demarcación; el rayo puede dirigirse oblicuamente. ¿Debiera tener el dueño de la tierra el derecho de prohibir tal penetración electrónica por considerarla un allanamiento? ¿Es relevante la posible aversión del dueño hacia el riesgo?

2. Si el gobierno subastara derechos de uso de las frecuencias radiofónicas, ¿sería la más elevada cantidad ofrecida por el postor igual al valor social esperado del uso? ¿Por qué podría no serlo?

3. ¿Debiera obligarse al gobierno a pagar por la información que obtiene de individuos y empresas para propósitos de los censos?

4. Supongamos que el gobierno expropia una propiedad en la que el inquilino ha hecho mejoras que sobrevivirán al arrendamiento. ¿Debiera incluirse el valor de la mejora en el monto de la compensación que el gobierno está obligado a pagar al inquilino por la expropiación? Véase Almota Farmers Elevator Warehouse Co. *vs.* United States, 409 U.S., 470 (1973). Y, hablando de la expropiación de propiedades en las que hay intereses divididos, ¿es preferible valuar cada interés y sumarlos todos, o valuar el título simple y dejar que los tenedores de intereses se dividan el monto? ¿Qué diferencia generaría esto, si acaso? ¿Y qué implican las respuestas de usted a estas interrogantes acerca de la adecuación o conveniencia de las disposiciones del arrendamiento que estipulan que, en el caso de una expropiación, recibirá el terrateniente toda la compensación? Véase Victor P. Goldberg, Thomas W. Merrill y Daniel Unumb, "Bargaining in the Shadow of Eminent Domain: Valuing and Apportioning Condemnation Awards between Landlord and Tenant", 34 *UCLA L. Rev.,* 1083 (1987).

5. Una cascada escénica corre por la propiedad de A. ¿Debiera permitirse que A se apropiara de un derecho a la cascada contra otros usuarios de la corriente (suponiendo que es el reclamante preferente)? ¿Hay algunas dificultades técnicas para el reconocimiento de tal derecho?

6. Si no hubiera ningún derecho para impedir que un hombre construyera de tal modo que obstruyera las ventanas de un vecino, ¿se estimularía así la extorsión?

7. ¿Son equivalentes las formulaciones siguientes del criterio económico para decidir a cuál de las partes deberá asignarse el derecho de propiedad? Si no lo son, ¿cuál es la mejor?

a) Asignar el derecho de tal modo que se minimicen los costos de transacción.
b) Asignar el derecho de tal modo que se maximice la probabilidad de que, si el derecho es más valioso para alguien diferente al tenedor del derecho, ocurra una transacción.
c) Asignar el derecho de tal modo que se vuelva innecesaria una transacción para alcanzar la eficiencia.
d) Asignar el derecho de tal modo que se minimice la suma de los costos de transacción y las pérdidas resultantes del hecho de que no se hagan intercambios maximizadores del valor.

8. Un hombre encuentra en su propiedad la entrada a una gran cueva escénica. La cueva se extiende también bajo la tierra de otras personas. ¿De-

biera otorgarse el derecho de propiedad de la cueva a este hombre, o debiera recibir cada uno de los propietarios superyacentes la porción de la cueva que se encuentre directamente debajo de su tierra, o debieran considerarse propietarios conjuntos e indivisos de la cueva? Véanse Edwards *vs.* Sims, S. W. 2d 619 (Ky., 1929); Richard A. Epstein, "Holdouts, Externalities, and the Single Owner: One More Salute to Ronald Coase", 36 *J. Law & Econ*., 553, 564-567 (1993).

9. Se construye un sanatorio para tuberculosos en un área residencial. Los valores de la propiedad declinan porque los residentes del área temen contagiarse con los pacientes del sanatorio. Su temor no tiene ninguna base científica. ¿Debiera considerarse, sin embargo, que el sanatorio es un perjuicio si la baja de los valores de la propiedad residencial es mayor que el aumento del valor del terreno usado para el sanatorio? Véase Everett *vs.* Paschall, 111 P. 879 (Wash., 1910); compárese el XXX.2, más adelante.

10. Supongamos que compra un colector solar caro y lo instala en su techo. Más tarde, el propietario de un terreno vecino construye un edificio que obstruye su colector. ¿Cuál recurso legal tendría en un sistema eficiente de derecho de la propiedad? Véanse Stephen F. Williams, "Solar Access and Property Rights: A Maverick Analysis", 11 *Conn. L. Rev.,* 430 (1979); Prah *vs.* Maretti, 321 N. W. 2d 182 (Wis., 1982).

11. Examinando la cuestión desde el punto de vista de la economía, ¿aprueba o reprueba los esfuerzos de las naciones que tienen costas por extender sus fronteras nacionales hasta 200 millas o unos 32 kilómetros de la costa? ¿Debiera distinguirse entre la afirmación de la autoridad nacional sobre el transporte marítimo y sobre la explotación de minerales?

12. Considere los hechos de Rhone *vs.* Stephens, [1994], 2 All E.R. 65 (H.L.). En 1960, el propietario de una casa y su cabaña anexa vendió el anexo, prometiendo en la escritura de venta que él y los propietarios sucesores de la casa mantendrían el techo común en beneficio del comprador del anexo y sus sucesores. En 1986, los propietarios actuales del anexo demandaron a los propietarios actuales de la casa a fin de obligarlos a reparar una gotera en la parte del techo sobre la cabaña. Perdieron. El tribunal dijo que si la promesa (convenio) contenida en la escritura hubiese sido negativa —por ejemplo, una promesa de no alterar el techo—, un incumplimiento de la promesa habría sido un intento de alargamiento de los derechos retenidos por el vendedor y se habría prohibido. Pero, dado que la promesa era positiva —el mantenimiento del techo en buen estado—, el cumplimento forzoso de la promesa habría impuesto una obligación contractual a alguien que no había firmado ningún contrato, o sea, al propietario actual de la casa. El propietario actual había comprado con conocimiento del convenio, porque se encontraba en la escritura original. ¿Puede imaginar cualquier razón por la que el propietario actual no debiera estar vinculado? (Hay algunas objeciones económicas contra el cumplimiento forzoso de los convenios restrictivos que no

"corresponden y conciernen" a la tierra: véase antes III.8, pero el convenio de reparación sí lo hacía.)

13. La Ley de Especies en Peligro de Extinción, 16 U.S. C. 1531, faculta al gobierno federal a designar hábitats de las especies en peligro. Una vez que se designa uno de tales hábitats, el propietario de la tierra que contiene ese entorno no podrá usar la tierra en una forma que ponga en peligro al hábitat. No hay ninguna compensación por la disminución del valor comercial de la tierra. ¿Tiene más sentido económico que el gobierno use su facultad de dominio eminente para adquirir estos hábitats antes que su poder regulador? Véanse Richard A. Epstein, "Babbitt *vs.* Sweet Home Chapters of Oregon: The Law and Economics of Habitat Preservation", 5 *Supreme Ct. Econ. Rev.*, 1 (1997); Dean Lueck y Jeffrey A. Michael, "Preemptive Habitat Destruction under the Endangered Species Act", 46 *J. Law and Econ.*, 27 (2003).

14. Supongamos que una persona hace mejoras por error en la tierra de otra persona (quizá piense que es suya, pero no se han satisfecho las condiciones necesarias para la obtención de un título mediante la prescripción adquisitiva). Compare dos remedios: *1)* el otorgamiento, al propietario verdadero, de la opción de conservar la tierra y pagar a la otra persona el valor de mercado de las mejoras; *2)* obligar a esa persona a comprar la tierra al propietario verdadero a su valor de mercado sin mejora. ¿Cuál remedio es más eficiente? ¿Y debiera importar que el error fuese intencional, negligente o puramente accidental? Véanse Thomas J. Miceli y C. F. Sirmans, "The Mistaken Improver Problem", 45 *J. Urban. Econ.*, 143 (1999); Gideon Parchomovsky y Alex Stein, "Reconceptualizing Trespass", 103 *Nw. U. L. Rev.*, 1823 (2009). *Cf.* después IV.15.

15. Supongamos que un cazador de tesoros puede obtener imágenes de un tesoro oculto mediante el uso de una cámara de video a control remoto. Si es el primero en descubrir el tesoro, ¿debiera permitírsele establecer un derecho de propiedad sobre él a causa de esas imágenes, antes de que haya hecho un esfuerzo por obtener su posesión? Véase Columbus-American Discover Group Inc. *vs*. Atlantic Mutual Ins. Co., 974 F.2d, 450 (4th Cir., 1992).

IV. DERECHOS Y REMEDIOS CONTRACTUALES

IV.1. El proceso del intercambio y los papeles económicos del derecho contractual

En el último capítulo destacamos la importancia de los intercambios voluntarios en la transferencia de recursos de usos menos valiosos a usos más valiosos, y señalamos diversos obstáculos para los intercambios maximizadores del valor; pero supusimos que el *proceso* de intercambio, una vez que se convienen los términos, opera confiablemente sin ninguna intervención de la ley.[1] Sin embargo, esto es estrictamente cierto sólo cuando ambas partes cumplen sus obligaciones de acuerdo con el contrato de manera simultánea. Cuando no se da la simultaneidad, surgen dos peligros para el proceso —el oportunismo y las contingencias imprevistas— que la ley trata de remediar.

A ocupa a B para que le construya una casa, con el pago a la terminación de la obra. Mientras se construye la casa y antes de que se haga cualquier pago, B se encuentra a merced de A, porque le resultaría difícil (¡especialmente si A es propietario del terreno en el que se está construyendo la casa!) vender la casa a cualquier otro si A decidiera no pagar por ella. Por lo tanto, en ausencia de un contrato legalmente exigible, A podría obligar a B a rebajar su precio una vez que la construcción se ha iniciado. (Dado que el derecho contractual, como cualquier otra institución social, no funciona perfectamente, no es sorprendente saber que los constructores insisten en recibir pagos progresivos, y no porque sus clientes sean sus prestamistas de costo menor.) Una vez que la casa ha sido construida y que A le paga a B, sus papeles se invierten. A se encuentra ahora a merced de B. Porque la construcción de la casa no es realmente el final del desempeño de B, sino el inicio. A está contando con recibir una corriente de servicios de la casa durante muchos años. Si B ha construido la casa defectuosamente, las expectativas de A se verán frustradas.

Al igual que el desincentivo para el cultivo de la tierra en un mundo sin derecho de propiedad, el problema del oportunismo contractual surge del carácter secuencial de la actividad económica. Si la siembra y la cosecha fuesen simultáneas, sería menos urgente la necesidad de reconocimiento de derechos de propiedad sobre la tierra (por oposición a los cultivos cosecha-

[1] Por lo que toca al derecho contractual, véanse E. Allan Farnsworth, *Farnsworth on Contracts* (3 vols., 3ª ed., 1998), y Alan Schwartz y Robert E. Scott, *Commercial Transactions: Principles and Policies* (2ª ed., 1991); este último es un libro de casos particulares con muchas notas económicas.

dos). Asimismo, si los intercambios contractuales fuesen simultáneos, sería menos urgente la necesidad de protección legal de los derechos contractuales.[2] Puesto que no lo son, la ausencia de derechos legalmente exigibles sesgaría la inversión hacia las actividades económicas que pudieran completarse en breve tiempo, y esto reduciría la eficiencia. A desea vender su vaca. Hay dos postores: B y C. La vaca vale 50 dólares para B y 100 para C (y sólo 30 para A), de modo que la eficiencia requiere que la vaca se venda a C antes que a B. Pero B tiene 50 dólares en la mano mientras que C no puede obtener nada de efectivo por una semana. C le promete pagar 75 dólares a A en una semana, y supongamos que esta prima de 25 dólares compensaría plenamente a A por los costos, en caso de incumplimiento, del juicio por daños o por la devolución de la vaca, descontados por el riesgo del incumplimiento, si la ley volviera exigible la promesa hecha por C a A. Pero si la ley no vuelve exigibles tales promesas, A podría decidir que, dado que en el intervalo C podría cambiar de opinión o B mientras tanto perder interés en la transacción, se encontrará en mejor situación vendiendo ahora la vaca a B. Si lo hace, el hecho de que la ley no provea un remedio si C incumple su promesa habrá inducido una mala asignación de recursos. B podría revender la vaca a C, para quien, recuérdese, vale más que para B, pero esto implicaría un costo de transacción adicional.

Supongamos ahora que D ofrece vender una camisa en cinco dólares y su competidor, E, ofrece una por seis dólares que, según afirma (con razón), durará tres veces lo que la camisa de D, de modo que tiene mayor valor. Si la diferencia en durabilidad no es evidente en una inspección o un manejo informales, E podría estar dispuesto a garantizar la mayor durabilidad de su camisa; pero si su promesa no es legalmente exigible, los consumidores podrían dudar de la honestidad de sus afirmaciones y comprar en su lugar la camisa de D, lo que de nuevo es un resultado subóptimo.

El sistema de intercambio voluntario no fallaría por completo sin un derecho contractual. Hay contratos en las sociedades que no tienen una maquinaria formal para el cumplimiento de la ley y, como veremos en el último capítulo de este libro, hay contratos entre naciones que no reconocen ninguna restricción legal a su soberanía. Algunas personas se dan a conocer por honrar sus promesas. Por ejemplo, el altruismo dentro de la familia (véase el capítulo siguiente) puede ser un garante de que la promesa se cumplirá. Como lo puede ser la reciprocidad, un sistema de intercambio basado en el reconocimiento del beneficio mutuo y en la inducción de la confianza.[3] Y, dado que

[2] Esto implica, y así lo vemos, que en las sociedades primitivas, en las que los intercambios contractuales tienden a ser simultáneos, el derecho contractual, en contraste con muchas otras áreas del derecho primitivo, es rudimentario; en particular, los contratos ejecutorios —aquellos en los que ninguna de las partes ha empezado todavía a cumplir su compromiso contractual— no se cumplen forzosamente. Richard A. Posner, *The Economics of Justice,* 182-184 (1981).

[3] Véase Benjamin Ho y David Huffman, "Trust and the Law", en *Search Handbook on Beha-*

alguien de quien se sabe que no cumple su parte de las negociaciones batallará para encontrar a alguien que desee hacer intercambios con él en el futuro, podría haber incluso una definición *más* explícita, por escrito o por incorporación explícita de la costumbre, de los compromisos de las partes de un intercambio que bajo un régimen de contratos exigibles. Las partes que negocian estarían particularmente deseosas de minimizar los malos entendidos que podrían causar acusaciones de mala fe, porque alguien contra quien se formularan tales acusaciones tendría mayores dificultades para encontrar personas que hicieran intercambios con él en el futuro.

Sin embargo, un sistema puramente voluntario de contratos no llevados a cumplimiento no sería eficiente. Aparte de los costos implicados en el mantenimiento de oficinas de crédito y la administración de depósitos de garantía (especialmente en un mundo donde no pudiera exigirse legalmente la devolución del depósito), no siempre funcionaría la autoprotección. Aunque alguien que considerara el incumplimiento de su contrato tomaría en cuenta los costos que tendría para él la disminución consiguiente de la disposición de otras personas a celebrar contratos con él en el futuro, los beneficios del incumplimiento podrían superar a tales costos. Esta persona podría ser muy anciana; o bien (algo relacionado) el contrato particular podría ser mucho mejor que todos los contratos que pudiera esperar en el futuro; o quizá no dependiera de su capacidad para hacer contratos, sino que sería capaz de funcionar excelentemente en el futuro sobre una base de tratos instantáneos en efectivo.

Así pues, la función fundamental del derecho contractual (reconocida como tal desde la época de Hobbes[4]) es, al disuadir a los individuos de un comportamiento oportunista en relación con sus contrapartes contractuales, alentar la coordinación oportuna y óptima de la actividad económica y (lo que es lo mismo) eliminar las costosas medidas de autoprotección. Pero no siempre es obvio cuándo se está comportando una de las partes de una manera oportunista. Supongamos que A contrata a B para que pinte su retrato "a satisfacción de A". B pinta un retrato que los conocedores admiran, aunque no lo suficiente para comprarlo ellos mismos al precio del contrato. A rechaza el contrato y se niega a dar razón alguna del rechazo. Si el rechazo no se hace de buena fe se considerará que A ha incumplido el contrato. El cumplimiento de buena fe —que en este contexto significa que no trata de aprovecharse de las vulnerabilidades creadas por el carácter secuencial del cumplimiento contractual— es un término implicado de todo contrato. Nadie se colocaría voluntariamente a merced de la otra parte, de modo que es razonable

vioral Law and Economics (Kathryn Zeiler y Joshua C. Teitelbaum [comps.], próxima publicación). Compárese la discusión anterior de la teoría de juegos en I.4.

[4] Thomas Hobbes, *Leviathan,* 70-71 (1914 [1651]) [edición en español: *Leviatán, o la materia, forma y poder de una república eclesiástica y civil,* 3ª. ed., trad. de Manuel Sánchez Sarto, FCE, México, 2017].

suponer que si las partes hubiesen pensado en la posibilidad de la mala fe lo hubiesen prohibido expresamente.

¿Debiera la ley ir más lejos y leer en el contrato una obligación implícita de ser razonable por parte de A? No debiera hacerlo (y no lo hace). Probablemente, las partes quisieron decir que A sería el único juez de la adecuación del cumplimiento de B. El lenguaje del contrato así lo sugiere, aunque no concluyentemente, y la sugerencia se refuerza al meditar sobre la incompetencia de un juez o un jurado para determinar si A, aunque de hecho (suponiendo que está actuando de buena fe) está insatisfecho con el retrato, estaba obligado a considerarlo de su agrado. En cambio, si el contrato, conteniendo el mismo lenguaje, se refiriera a la pintura de las paredes exteriores de una fábrica, el tribunal podría decidir que las partes no habían intentado hacer del capricho del comprador la medida del cumplimiento del vendedor, ya que el juez o el jurado podrían determinar sin gran dificultad si el trabajo de pintura era adecuado para su propósito ordinario.[5]

Cambiemos el ejemplo. A, un fabricante, otorga a B, un comerciante, una representación exclusiva para cierta área. Es decir, A conviene en no vender su producto a nadie más en esa área durante el periodo del contrato. En una opinión famosa del juez Benjamin Cardozo, el tribunal sostuvo que un contrato de representación exclusiva contiene una condición implícita de que el representante hará sus mejores esfuerzos para vender el producto del proveedor.[6] Sin tal condición, B podría privar al contrato de todo valor para A simplemente vendiendo en cambio los productos de otros fabricantes.

Este ejemplo revela que otro nombre para el oportunismo es el de monopolio. El contrato entre el fabricante y el distribuidor dio a este último un monopolio porque privó a aquél del derecho de vender a distribuidores rivales dentro del área especificada en el contrato. La ley supone que las partes no preveían que el distribuidor estuviera en libertad para aprovecharse del monopolio, de modo que interpola una condición de los mejores esfuerzos. Es lo que se llama una cláusula de incumplimiento; las partes pueden rechazarla o modificarla, pero, si no lo hacen, ésta rige su contrato a pesar de que no es un término explícito del mismo.

Podría argüirse que si el fabricante hubiera deseado tal protección la habría negociado, pero tal vez no lo haya hecho porque las partes preferían evitar la posibilidad de un litigio por el significado de los "mejores esfuerzos". Pueden haber preferido depender en lugar de ello del interés del distribuidor en contratos futuros, o de la brevedad del término del contrato, o de la inclusión de una cláusula que permitiera a ambas partes dar por terminado el contrato con un aviso a corto plazo, a fin de proteger al fabricante contra la posibilidad de que el distribuidor explotara el monopolio conferido por el contrato.

[5] Morin Building Products Co. *vs.* Baystone Construction, Inc., 717 F.2d, 413 (7th Cir., 1983).
[6] Wood *vs.* Duff-Gordon, 118 N.E. 214 (N.Y., 1917).

En otras palabras, algunas omisiones de la protección contractual podrían ser deliberadas: el producto del sacrificio realizado a cambio de obtener una ventaja dentro del ámbito que se extiende desde el peligro del oportunismo, por una parte, hasta los costos directos e indirectos del litigio (incluido el riesgo del error), por la otra.[7] Sin embargo, otras omisiones se pueden atribuir a las limitaciones de la previsión. Entre más largo sea el periodo del cumplimiento de un contrato—y recuérdese que el "cumplimiento" incluye toda la corriente de servicios futuros que se contempla en el intercambio—, más difícil será que las partes prevean las diversas contingencias que podrían afectar al cumplimiento. Las personas que consideran una transacción pueden reducir los costos de los errores potenciales derivados de la previsión imperfecta acortando la duración de su contrato (considérese el empleo a voluntad, o los mercados de efectivo), ya que el futuro próximo es más predecible que el futuro lejano. O, lo que es lo mismo, pueden convenir sólo en algunas cosas y dejar el resto para una negociación futura ("aceptando un acuerdo"). Otra alternativa es la sustitución de la contratación por la integración vertical: por ejemplo, un productor puede optar por elaborar en lugar de comprar un insumo, aunque en realidad esto no sería más que la sustitución de un contrato para comprar el insumo por los contratos de trabajo, además de los contratos para comprar cualquier material físico que se utilizaría en la fabricación del insumo en la fábrica. Las cláusulas implícitas, como la cláusula de los mejores esfuerzos, son otra alternativa (aquella con los que comenzamos).

Otra alternativa es que un tribunal o un árbitro llenen cualquier vacío en el contrato si y cuando la omisión sale a relucir y origina una disputa. Tal compleción de omisiones adjudicativa adecuada es un método particularmente atractivo de resolver las contingencias que, incluso si ambas partes están plenamente conscientes de que *podrían* ocurrir, es poco probable que lo hagan. Porque los costos de una redacción cuidadosa para resolver las contingencias podrían superar a los beneficios, cuando tales beneficios se descuentan por la poca probabilidad de que realmente se materialicen. Puede resultar más barato, desde el punto de vista social, que el tribunal redacte (después de los hechos, por supuesto, una vez surgido un caso) el término contractual necesario para afrontar la contingencia imprevista y no considerada cuando ocurra. Un fallo podría equivaler a mil sesiones de redacción. Y, de esta manera, si un tribunal está convencido de que la mayoría de las partes en los contratos de negociación exclusiva estarían de acuerdo en imponer al distribuidor una obligación del mejor esfuerzo, entonces darle importancia a tal obligación en cada contrato de negociación exclusiva que no anula

[7] Este tema se discute con mayor detalle en XII.17. Véanse Douglas G. Baird, "Self-Interest and Corporation in Long-Term Contracts", 19 *J. Legal Stud.,* 583 (1990); Alan Schwartz, "Relational Contracts in the Courts: An Analysis of Incomplete Agreements and Judicial Strategies", 21 *J. Legal Stud*., 271 (1992).

tal obligación evita costos de negociación y redacción. La interpolación judicial de la disposición es eficiente si las probabilidades de que las partes se excluyan son menores que las de participar (si no fuera una cláusula de incumplimiento).

La observación del comportamiento de exclusión voluntaria puede proporcionar una prueba empírica de la eficacia de una cláusula de incumplimiento. Es probable que una disposición de la que rara vez se excluyen las partes sea eficiente; pero, si se excluyen de manera muy frecuente, es probable que sea ineficiente. Sin embargo, los estudios han encontrado que las reglas del derecho contractual no dan lugar con frecuencia a la exclusión de las partes aun si son ineficientes.[8] Los autores de estos estudios atribuyen este resultado al "sesgo del *statu quo*", interpretado en términos psicológicos. Con todo, una explicación alternativa es que, dado que muy pocas disputas contractuales dan lugar a un litigio, los costos de transacción de la negociación de una modificación en la disposición estándar excederán con frecuencia a los beneficios incluso si la disposición es ineficiente. Una razón es que, en la negociación de un contrato, la parte que presiona por una modificación en una disposición estándar puede despertar sospechas en la otra parte de que se está anticipando al litigio y está buscando tomar ventaja en él.

No obstante, si las partes contratantes realmente prefirieran hacerlo sin la cláusula legalmente exigible de los mejores esfuerzos a fin de reducir el riesgo del litigio, cada vez más contratos contendrían disposiciones que renunciaran a cualquier obligación de utilizar los mejores esfuerzos; el litigio de los mejores esfuerzos desaparecería y finalmente se modificaría la regla de incumplimiento. Nada de esto ha sucedido.

En ocasiones los costos de la compleción judicial de omisiones serán prohibitivos, y entonces el tribunal se niega a completar el vacío, como en la negativa del derecho común a exigir el cumplimiento de un contrato que carece de un término de precio o de cantidad. La alternativa de la interpolación de una cláusula de "precio razonable" o "cantidad razonable" se rechaza porque sería demasiado oneroso para un tribunal calcular qué precio o cantidad habrían negociado las partes. Sin embargo, de acuerdo con la razón detrás de la regla, el Código Comercial Uniforme (vigente en todos los estados de la Unión Americana) permite que las partes omitan el precio si hay un precio al contado publicado. No sólo la determinación de los precios en estas circunstancias es mecánica, sino que la especificación del precio en el contrato frustraría el propósito de éste que es negociar a un precio de mercado en constante cambio y no a un precio fijo.

[8] Omri Ben-Shahar y John A. E. Pottow, "On the Stickiness of Default Rules", 33 *Fla. St. U. L. J.*, 651, 654 (2006); Russell Korobkin, "The Status Quo Bias and Contract Default Rules", 83 *Cornell L. Rev.*, 608 (1998). Véase también Lisa Bernstein, "Social Norms and Default Rules Analysis", 3 *S. Cal. Interdisciplinary L. J.*, 59, 71-72 (1993).

El llenado de las omisiones potenciales en los contratos debe distinguirse de la desambiguación de términos específicos, lo que conforma el centro del problema de la interpretación del contrato. Es cierto que tanto el llenado de las omisiones como la desambiguación son "interpretativos" en el sentido de que están dirigidos a dilucidar la manera en que las partes habrían resuelto el problema que surgió si lo hubieran previsto cuando negociaron su contrato. Pero, en el caso de ambigüedad, el tribunal no puede simplemente sacar una cláusula preconcebida del estante e introducirla en el contrato, seguro de que si la regla no fuera la adecuada las partes la habrían anulado explícitamente en su contrato.

¿Qué hacer con los casos en que las intenciones de las partes, tal como se adivinan por el lenguaje del contrato o por testimonios, divergen de la noción que tiene el tribunal de lo que sería el término eficiente que se debe interpolar en el contrato? ¿Debieran gobernar la eficiencia o las intenciones? La respuesta es las últimas. Los individuos que hacen una transacción son de ordinario jueces de su propio interés más confiables que un juez (o un jurado), quien no tiene ningún interés personal en el asunto en el que se embarcaron las partes al firmar el contrato ni posee algún conocimiento de primera mano al respecto. Pero las discrepancias existentes entre el acuerdo (aparente) y la eficiencia no deben ignorarse; pueden ser indicios importantes de la existencia de error, incapacidad u otras razones para creer que una interpretación literal del contrato no promueve los fines conjuntos de las partes.

Supongamos que A le compra bienes a B, para ser entregados en un mes, y durante ese mes se incendia el almacén de B y los bienes quedan destruidos. El contrato no dice nada acerca de la asignación del riesgo de la pérdida antes de la entrega. Pero dado que B puede prevenir un incendio (o asegurarse contra esa eventualidad) en su propio almacén a un costo menor que el de A, si las partes hubieran pensado sobre ese asunto habrían asignado probablemente el riesgo a B, aunque ya no sea el "propietario" de los bienes, y ésa es la asignación que el tribunal debería hacer en ausencia de prueba en contrario acerca de las intenciones de las partes.

Generalizando a partir de este ejemplo, podemos esperar que el fabricante garantice las dimensiones del cumplimiento que se encuentren primordialmente bajo su control y no bajo el control del comprador. Por lo tanto, se supone que el fabricante debe garantizar que los bienes son adecuados para el uso al que se destinan, pero no que durarán indefinidamente; su durabilidad podría depender de la forma en que use los bienes el comprador. El estado de Wisconsin contrató en cierta ocasión a un hombre llamado Bentley para que construyera aleros en el capitolio estatal bajo la dirección del arquitecto del estado. Bentley siguió los planos del arquitecto al pie de la letra, pero los planos no eran buenos y los aleros se derrumbaron poco tiempo después de su terminación. El estado demandó a Bentley alegando que había garantizado su trabajo contra tal calamidad. El contrato no decía nada apli-

cable a ese asunto; obviamente, ninguna de las partes había creído probable que los aleros se derrumbaran porque los planos del arquitecto fueran malos. El estado perdió este juicio.[9] Éste es el resultado económico correcto. El estado pudo haber prevenido la calamidad a un costo menor que el de Bentley, mediante una selección o una supervisión más cuidadosa del arquitecto. Aun así, ¿no podría haber sido uno de los propósitos del contrato el aseguramiento del estado contra el colapso de los aleros por cualquier causa? El aseguramiento es un procedimiento para afrontar contingencias imprevistas, y los contratos son a menudo un método de aseguramiento (véase después IV.5). Pero es improbable que Bentley fuese un mejor asegurador que el estado. Bentley tendría que haber ido a comprar una póliza de seguros; el estado podía autoasegurarse contra el riesgo particular.

Los ejemplos discutidos hasta ahora ilustran el punto importante de que, dada la libertad contractual, el derecho contractual no puede utilizarse fácilmente para lograr objetivos distintos de la eficiencia: Un fallo judicial para descubrir la solución eficiente de una disputa contractual puede rectificarse para el futuro mediante la modificación de la redacción.

Hasta ahora hemos recurrido a la noción del derecho contractual como un auxiliar del intercambio, pero ésta es una visión demasiado estrecha del asunto, como lo ilustrarán algunos ejemplos:

1. Un hombre rico promete en un momento de euforia pagar toda mi instrucción universitaria. Renuncio a mi empleo de medio tiempo, pero el hombre rico rompe después su promesa, y no puedo conseguir un nuevo empleo.
2. A promete entregarle a B ciertos bienes "el día 12". B cree que A se refería al 12 de este mes, pero en realidad A quiso decir el 12 del mes siguiente: simplemente no podía entregar los bienes tan pronto como esperaba B (sin que lo supiera A).
3. Una compañía siderúrgica se compromete a entregar acero a una compañía constructora dentro de 60 días, pero la compañía siderúrgica está cerrada por una huelga ilegal y no puede hacer la entrega dentro de ese plazo.

En ninguno de estos casos se trata de saber si una de las partes de un intercambio ha *rehusado* cumplir su parte de la negociación. En el primer caso no hay ningún intercambio; la renuncia a mi empleo de medio tiempo no confiere ninguna ventaja al rico promitente. Es posible que él ni siquiera haya sabido que yo renuncié a mi empleo. En el segundo caso tampoco hay intercambio; las partes intentaban transacciones diferentes. En el tercer caso, el cumplimiento se hizo imposible por circunstancias que escapaban al

[9] Bentley *vs.* State, 41 N. W. 338 (Wis., 1889).

control del promitente. Pero en los tres casos hay un argumento económico a favor de la imposición de sanciones a la parte que incumplió y, por lo tanto, a favor de proponer un contrato. La promesa ociosa del hombre rico indujo una confianza que le costó mucho al promisario cuando la promesa se rompió, y tal costo puede evitarse para el futuro considerando responsable a un promitente por el costo de la confianza del promisario. Supongamos, en el segundo caso, en el que el comprador y el vendedor confunden la fecha, que la costumbre en la industria es que una fecha de entrega sin especificación del mes se refiere al mes en curso. A es nuevo en la industria e ignora la costumbre. Sin embargo, si se le hace responsable de la promesa tal como la entendió B, se obtendrá el efecto saludable de inducir a los recién llegados a dominar pronto el lenguaje de la industria… aunque para confiar en que éste sería el resultado óptimo tendríamos que considerar: *1)* si las empresas existentes no podrían ser la fuente más barata de la información de los recién llegados acerca de la costumbre, y *2)* los posibles efectos anticompetitivos de echar la carga de la adquisición de esta información sobre los hombros de los recién llegados.[10]

En cuanto al tercer caso, la compañía siderúrgica se encuentra probablemente en mejor posición que el comprador para prever y tomar salvaguardias apropiadas contra una interrupción de la producción causada por una huelga ilegal. En tal caso, la asignación del riesgo de tal interrupción a la compañía siderúrgica, considerándola responsable de los daños causados por la demora al comprador, podría ser el procedimiento más barato para la minimización de los costos de tales demoras en el futuro.

En resumen (y como un anticipo), el derecho contractual tiene cinco funciones económicas claras: *1)* prevenir el oportunismo, *2)* interpolar términos eficientes, ya sea sobre una base mayorista o una base minorista (compleción de omisiones *versus* interpretación adecuada), *3)* sancionar los errores evitables en el proceso de contratación, *4)* asignar el riesgo a quien pueda soportarlo mejor, y *5)* reducir los costos de la resolución de las disputas contractuales.

IV.2. La consideración

La doctrina de que una promesa, para ser legalmente exigible, debe estar apoyada por la consideración puede parecer a primera vista un corolario lógico de la idea de que el papel del derecho contractual es la facilitación, mediante el intercambio voluntario, del movimiento de los recursos a sus usos más valiosos. Si la promesa es enteramente unilateral no puede formar parte del proceso de intercambio. Pero no es cierto que las únicas promesas

[10] Elizabeth Warren, "Trade Usage and Parties in the Trade: An Economic Rationale for an Inflexible Rule", 42 *U. Pitt. L. Rev*., 515 (1981).

que vale la pena hacer cumplir son las instrumentales para un intercambio. En el caso en que un hombre joven renuncia a un empleo de medio tiempo sobre la base de la promesa de un hombre rico de pagar toda su instrucción universitaria, no hay intercambio en ningún sentido realista (aunque haya un beneficio mutuo), pero la ley trata el incumplimiento de una promesa como un incumplimiento de contrato. Lo hace mediante la ficción de que la "confianza detrimental" (renuncia al empleo de medio tiempo) es una forma de la consideración, pero lo hace de todos modos.

Veamos un caso más difícil: A rescata a B de cierto peligro, y B le promete a A una anualidad por el resto de su vida. ¿Exigible? Sí, según la teoría (si se puede llamar así) de que hay una "consideración moral", aunque el rescate se había completado antes de que se hiciera la promesa. El resultado tiene sentido, siempre que B intente hacer una promesa legalmente vinculante. Al volver vinculante la promesa de B se transmitirá una información útil para A: que puede contar realmente con este dinero por el resto de su vida. Esta información vuelve el regalo más valioso para A y, por lo tanto, también para B, ya que debe obtener satisfacción de la satisfacción de A, o no haría el regalo.

Este ejemplo destaca el hecho paradójico de que a menudo podemos aumentar nuestra utilidad reduciendo nuestra libertad. Esta observación se encuentra detrás del derecho contractual. A está en mejor situación, no peor, cuando hace una promesa vinculante de pagarle a B para que le construya una casa, porque si no estuviera en posibilidad de dar una seguridad de pago A podría tener que pagar por adelantado el precio completo, o depositar con algún mediador una propiedad de valor igual al precio de la casa, o quizá prescindir de la casa por entero. El verdadero misterio en los casos de "consideración moral" es por qué la ley no crea simplemente una *forma* de hacer promesas vinculantes sin requerir la consideración o establecer clases de casos especiales en los que no se requiera la consideración. La ley solía hacer esto. Las promesas hechas bajo sello eran exigibles sin consideración. Ése era un instrumento útil; su desaparición es un enigma.

Éstas son algunas de las funciones económicas que podríamos creer que desempeña el requerimiento de una consideración:

1. Reduce el número de juicios contractuales fraudulentos requiriendo que el demandante pruebe algo más que simplemente que alguien le prometió algo y que incumplió; el demandante debe demostrar que hubo un trato de alguna clase, lo que resulta un poco más difícil que fabricar una ficción de punta a punta. Esta función probatoria es importante en un sistema como el nuestro que hace cumplir los contratos orales: una virtual necesidad en la era del teléfono, pero quizá no en la era del *e-mail*.
2. Reduce la probabilidad de los compromisos contractuales inadvertidos por efecto de un uso descuidado, precipitado o casual del lenguaje promisorio. Al respecto, la consideración se asemeja a los periodos de

"enfriamiento", lo que permite a un consumidor rescindir un contrato sin sanción en un plazo corto y fijo, que la ley de protección de los consumidores impone en ocasiones. Éstos son recursos relativamente baratos para proteger a la gente contra las deficiencias cognitivas comunes.[11] Sin embargo, adviértase, por un lado, la tensión entre los puntos 1 y 2 y la doctrina de la consideración moral, por el otro.

3. Liberar a los tribunales de tener que hacer cumplir un montón de promesas triviales que se hacen en ambientes sociales y familiares. El concepto de la confianza detrimental confina el cumplimiento forzoso a las promesas gratuitas que probablemente induzcan a los promisarios a incurrir en costos sustanciales. Una razón mejor para negarse a exigir el cumplimiento de esas promesas (puesto que las demandas por cantidades pequeñas de dinero son poco frecuentes) es evitar la implicación judicial en disputas personales que se resuelven mejor de manera informal que en el tribunal.
4. Alejar de los tribunales los casos en que, aunque el contexto es de intercambio, el compromiso de una de las partes, o de ambas, se expresa de modo vago, como en el ejemplo que discutimos en la sección anterior de este capítulo de la omisión del término de precio del contrato.
5. Disuasión del comportamiento oportunista. En Alaska Packers' Assn. *vs*. Domenico,[12] el demandado contrató marineros en San Francisco para un viaje a Alaska a fin de pescar salmón. Pero, cuando el barco del demandado llegó a las aguas de Alaska, los marineros anunciaron que no trabajarían si el demandado no aumentaba los salarios convenidos antes del inicio del viaje. Incapaz de obtener mano de obra sustituta durante la corta estación de la pesca en Alaska, el demandado se rindió a la amenaza (una alternativa habría sido una contraamenaza: suspender el viaje y privar así a los marineros de cualquier salario; esto sería análogo a un cierre patronal en un conflicto laboral común; véase el capítulo XII) y prometió pagar a los marineros los salarios mayores que demandaban cuando el barco regresara a San Francisco. El demandado se retractó cuando el barco regresó a San Francisco, y los marineros demandaron… y perdieron el juicio, porque se consideró que la modificación del contrato no era legalmente exigible en virtud de que no estaba apoyada por una consideración nueva. Éste es el resultado económico correcto, porque, una vez que se conozca que tales modificaciones no son legalmente exigibles, los trabajadores que se encuentren en la posición de los marineros en el caso Domenico sabrán que de nada les servirá aprovecharse de la vulnerabilidad de sus empleadores.[13]

[11] Véase, por ejemplo, Cass R. Sunstein, "Boundedly Rational Borrowing", 73 *U. Chi. L. Rev.*, 249 (2006).

[12] 117 F. 99 (9th Cir., 1902).

[13] Véanse otros análisis económicos de los casos de Alaska Packers y Goebel en Varouj A. Aiva-

En Goebel *vs.* Linn se había llegado al resultado contrario.[14] El demandante había prometido abastecer de hielo al demandado, pero la "cosecha" de hielo fracasó debido al tiempo atípicamente cálido, y el proveedor del hielo exigió y obtuvo la promesa del demandado de que le pagaría un precio mayor. El demandado se retractó, y a la acción de incumplimiento del contrato ejercitada por el demandante se opuso la excepción de que la modificación no se había apoyado en ninguna consideración. Pero las pruebas demostraron que el demandante (muy al revés de los marineros en Domenico) se habría ido a la quiebra si el demandado hubiese tratado de hacer cumplir el contrato tal como se había escrito, y si ello hubiera ocurrido el demandado podría haberse quedado sin hielo en absoluto. La modificación no había sido oportunista de parte del demandante —no había ninguna sugerencia de que el demandado hubiese creado deliberadamente el riesgo de la quiebra—, sino que se trataba de un ajuste razonable ante un cambio imprevisto de circunstancias. Si no se hubiera modificado el contrato, en realidad el promitente no habría cumplido, mientras que en un caso oportunista, dado que el promitente puede cumplir con ganancia respetando los términos originales del contrato, es probable que cumpla si la ley priva de eficacia a su amenaza.

Consistentes con la interpretación económica de la doctrina de la consideración, los tribunales sólo investigan la existencia, no la adecuación, de una consideración. Preguntarse si hay una consideración es simplemente averiguar si la negociación ha llegado a un convenio. Ir más allá y preguntarse si la consideración es adecuada requeriría que el tribunal hiciera lo que está menos preparado para hacer que las partes: decidir si el precio y otros términos esenciales especificados en el contrato son razonables. Esto revela también la sabiduría de la tendencia judicial hacia el empleo de la doctrina de la coacción antes que hacia la doctrina de la consideración para decidir si hace cumplir una modificación contractual. Si los marineros del caso Domenico hubiesen sido astutos, habrían ofrecido trabajar media hora extra cada día a cambio del salario mayor. Bien, pero ¿qué diremos si hubiesen ofrecido trabajar dos segundos extras al mes? ¿En qué punto se vuelve la consideración tan nominal que ya no es ninguna consideración? La respuesta es que la inadecuación de la consideración es siempre potencialmente relevante como una prueba circunstancial de la coacción, el error, el fraude o alguna otra razón para descartar un contrato. Entre menos adecuada sea la consideración, más fuerte será el efecto de la prueba.

zian, Michael J. Trebilcock y Michael Penny, "The Law of Contract Modifications: The Uncertain Quest for a Bench Mark of Enforceability", 22 *Osgoode Hall L. J.*, 173 (1984).

[14] 11 N. W. 284 (Mich., 1882).

IV.3. Problemas de formación; el origen del asentimiento mutuo y los contratos unilaterales

La presunción de que un intercambio que se está considerando maximiza el valor sólo es válida cuando las partes convienen efectivamente en sus términos. Si usted ofrece comprar mi reloj en 10 dólares, pero la compañía de telégrafos comete un error de transmisión y el telegrama que recibo dice 20, el hecho de que yo acepte la oferta tal como yo la entiendo no es prueba de que la venta aumentará el valor; el reloj podría valer 14 dólares para mí y sólo 12 para usted. La teoría "subjetiva" del contrato, según la cual debe haber una conjunción efectiva de las mentes de las partes contratantes para que surja un contrato exigible, tiene así un sentido económico. Pero no se sigue de aquí que nunca se condenará al pago de daños a quien se niegue a cumplir una promesa que no habría hecho si no hubiera habido una falla en la comunicación: también la teoría objetiva del contrato tiene una base de justificación económica, aunque desde un punto de vista económico tal negativa se parece más a un cuasidelito que a un incumplimiento de contrato. Está diseñada para desalentar el comportamiento ineficiente antes que para implementar un acuerdo.

En el caso del telégrafo, la cuestión que interesa principalmente al economista consiste en saber cuál de las partes se encuentra en mejor posición para impedir los malos entendidos a resultas de una transmisión errada. A menudo se trata de la parte que selecciona el método de comunicación, que podría enviar una carta de confirmación o emplear el teléfono o un mensajero. Si esta parte pudo haber evitado el malentendido a un costo menor que el de la otra parte, entonces la asignación de la responsabilidad en su contra volverá menos probables los errores futuros. Pero cuando se asigna la responsabilidad por esta razón resulta engañoso decir que la comunicación defectuosa creó un contrato. El defecto imposibilita saber si se intentaba hacer un intercambio; la ley trata la comunicación abortada como un contrato a fin de evitar los errores futuros.

El contrato unilateral plantea un problema interesante de consentimiento mutuo. Ofrezco 20 dólares por la devolución del gato que se me perdió. No hay ninguna negociación con halladores potenciales, ninguna aceptación de mi oferta en el sentido convencional. Pero alguien que se entera de la recompensa y devuelve a mi gato tiene un derecho a la recompensa legalmente exigible; su cumplimiento de los términos de la oferta se trata como una aceptación. Este resultado es correcto porque promueve una transacción maximizadora del valor. El gato vale más que 20 dólares para mí y menos de 20 dólares para quien lo encontró, de modo que el intercambio de dinero por el gato aumenta el bienestar social. Adviértase que el enfoque del contrato unilateral economiza en el caso del gato perdido al evitar los costos de tener que negociar con antelación con todos los halladores potenciales.

Una cuestión más difícil consiste en saber si, como sostienen la mayoría de los tribunales, la persona que lo encontró debe saber de la oferta a fin de tener un derecho a la recompensa. La razón jurídica para contestar este interrogante afirmativamente, formal más que práctica, es que la aceptación requiere el conocimiento de la oferta. El problema económico consiste en saber si el requerimiento de la prueba del conocimiento se justifica por el costo en un caso en el que se ofrece una recompensa. Esto depende de que tal requisito aliente o desaliente la devolución del bien perdido. Para dar respuesta a esta cuestión debemos distinguir entre la persona que lo encuentra y que fue inducida a buscar el bien perdido por el conocimiento de que hay una recompensa por su devolución y el hallador casual, que no invertirá en la búsqueda, pero podría devolver el bien perdido si casualmente se lo encuentra, sabiendo que *podría* haber una recompensa. Una regla que requiera el conocimiento desalienta a la persona que lo encuentra casualmente, pero alienta al buscador activo al disminuir la probabilidad de que su competidor, el hallador casual, devuelva el bien si lo encuentra primero. Por lo que debemos considerar si el aumento de las devoluciones de los buscadores activos tenderá a ser mayor o menor que la disminución de las devoluciones de los halladores casuales. Los buscadores activos son menos comunes que casuales, pero ¿cuántos halladores casuales —quienes por definición no saben si hay o no una remuneración, sino sólo que podría haberla— se molestarán efectivamente por devolver lo que encontraron? El otorgamiento, a estos halladores, de un derecho legal a cualquier recompensa que se ofrezca podría no aumentar significativamente el número de las devoluciones por su cuenta; pero por la misma razón no desalentará tampoco grandemente a los buscadores activos. Es posible que ninguna de las dos reglas conduzca a más devoluciones, en cuyo caso la regla que requiere el conocimiento, al reducir el número de las reclamaciones legales, es preferible porque es más barata.[15] Es decir, más barata para los tribunales, pero también más barata para los dueños de los bienes perdidos, porque no tendrán que compensar al hallador casual que les devuelva la propiedad.

¿Cuándo, si acaso, debiera tratarse el silencio como la aceptación de una oferta? La respuesta obvia es que nunca. De otro modo, un vendedor podría bombardear a los consumidores con ofertas que establecieran que la oferta se consideraría aceptada a menos que el consumidor enviara por correo su rechazo antes de cierta fecha. Los consumidores tendrían que soportar un costo de correo cada vez mayor para protegerse contra la inundación de productos no deseados. Pero el análisis económico sugiere que la palabra "nunca" debiera cambiarse por la frase "algunas veces".[16] Dejando de lado la cuestión

[15] Esto supone que la reducción del número de reclamaciones no se ve contrarrestada por el aumento de la complejidad del procesamiento de cada reclamación. El aumento de la complejidad se debe a la presencia de una interrogante adicional: si el hallador sabía de la recompensa.

[16] Véase un análisis formal en Avery Katz, "The Strategy Structure of Offer and Acceptance: Game Theory and the Law of Contract Formation", 89 *Mich. L. Rev.,* 215, 249-269 (1990).

de los errores postales, debemos comparar los costos del correo bajo los dos regímenes: uno en el que el silencio se considere un rechazo; otro en el que se considere una aceptación. En el primer régimen, todo contrato aceptado requiere dos cartas (u otra forma de comunicación escrita) —la oferta por correo y una aceptación por correo—, pero todo contrato rechazado requiere sólo una carta. En el segundo régimen, todo contrato aceptado requiere sólo una carta, pero todo contrato rechazado requiere dos cartas. Los costos relativos de los dos regímenes dependerán de la razón de los contratos aceptados a los rechazados. Si tal razón es muy elevada, la regla de que el silencio es aceptación reducirá los costos del correo. Por lo tanto, esperaríamos que la ley tratara de identificar los contextos en los que puede esperarse una razón elevada de contratos aceptados a rechazados y aplicar la regla de que el silencio es aceptación en tales contextos, y la regla de que el silencio es rechazo en los otros. Y eso es lo que ocurre. Los tribunales se preguntan si el oferente, de ordinario debido a una relación de tratos anteriores con el ofertado, supone razonablemente que es muy probable que el ofertado acepte la oferta; en tal caso, se permitirá que el oferente especifique que el silencio será tratado como aceptación.

La regla del "apartado postal" del derecho contractual considera efectiva la aceptación cuando la envía el ofertado, no cuando la recibe el oferente. Esta regla permite que el ofertado empiece el cumplimiento (o las medidas preparatorias) inmediatamente después de enviar su aceptación, sin demorar el cumplimiento del oferente, el que en todo caso no se iniciaría antes de que éste recibiera la aceptación, ya que hasta entonces no sabría si había o no un contrato.

Por último, una licencia de "envoltura" es el resultado de la aceptación de una oferta en línea de un producto de software si la oferta requiere que el comprador acepte los términos del contrato detallados que no se le darán a conocer hasta después de que él indique su aceptación. En ProCD, Inc. *vs.* Zeidenberg,[17] el tribunal argumentó que esta aceptación era eficaz para crear un contrato, aunque el comprador no conociera los términos de la oferta cuando se comprometió a la compra. El receptor de la oferta no está obligado a aceptar; si lo hace apuesta a que el interés del oferente en la buena voluntad de los consumidores desalentará la inclusión de términos excesivamente onerosos.

IV.4. El error mutuo

En el famoso caso de la venta de la vaca Rose 2d of Aberlone,[18] tanto el vendedor como el comprador creían que la vaca era estéril, y el precio se fijó correspondientemente. De hecho, la vaca estaba preñada y valía cerca de 10

[17] 86 F.3d, 1447 (7th Cir., 1996).
[18] Sherwood *vs.* Walker, 66 Mich. 568, 33 N. W. 919 (1887).

veces más que el precio de venta. El error se descubrió antes de que la vaca fuera entregada al comprador, y el vendedor canceló la venta. El tribunal sostuvo la cancelación. Si aceptamos la versión de los hechos presentados en la opinión mayoritaria, el resultado podría parecer consistente con la eficiencia. No había ninguna base para presumir que la vaca valiera más en posesión del comprador que en posesión del vendedor —dado que su valor verdadero era de un orden de magnitud diferente del que las partes habían pensado—, y el vendedor no había sido descuidado al pensar que la vaca era estéril. El vendedor había cometido un error razonable; más aún, un error inevitable a un costo razonable. No había ninguna base para pensar que el cumplimiento forzoso del contrato maximizaría el valor (es decir, que la vaca más el ternero más la capacidad de producir terneros en el futuro valían más para el comprador que para el vendedor), y, dado que el vendedor no había sido descuidado, no había razón para "castigarlo" haciendo cumplir el contrato.

Pero el caso puede ser enfocado diferentemente observando que había ocurrido una contingencia imprevista y preguntando cómo habrían asignado las partes el riesgo de esta ocurrencia si la hubiesen previsto. Había algunas pruebas de que el precio de venta de Rose incluía su valor si estuviera preñada, descontado drásticamente por la probabilidad de esta feliz eventualidad. Las pruebas sugerían que las partes habían intentado transferir al comprador el riesgo de que la vaca resultara preñada, en cuyo caso el contrato debía haberse cumplido forzosamente. Pero aun sin tales pruebas hay un argumento a favor de asignar al vendedor el riesgo de que la vaca no sea lo que parece. En general, aunque no en todos los casos particulares, el propietario tendrá acceso, a un costo menor que el del comprador, a información acerca de las características de su propiedad, de modo que puede evitar errores acerca de estas características a un costo menor que el de los compradores potenciales. Por esta razón el vendedor de una casa es responsable ante el comprador por los defectos ocultos (por oposición a los obvios); podría emplearse un principio similar para decidir otros casos de error mutuo acerca de las características del producto en una venta.[19]

Cuando ocurre una contingencia que afecta el cumplimiento *después* de que el contrato ha sido firmado (Rose estaba preñada cuando se firmó el contrato), los tribunales tienen menor dificultad para reconocer que el problema consiste en saber cómo asignaron las partes implícitamente el riesgo de un acontecimiento inesperado. Al nivel más elemental, si un contrato establece la entrega de trigo en una fecha fija a un dólar la libra, el hecho de que el precio de mercado de ese día sea de dos dólares la libra no invalidará el contrato; las partes intentaban claramente asignar al proveedor el riesgo de los cambios de precio. Pero a veces la asignación calculada de los

[19] Véase un análisis contrario de Sherwood en Janet Kiholm Smith y Richard L. Smith, "Contract Law, Mutual Mistake, and Incentives to Produce and Disclose Information", 19 *J. Legal Stud.*, 467 (1990).

riesgos no está clara. Éste es el dominio de las doctrinas de la imposibilidad, la impracticabilidad, la frustración y la *force majeure,* que discutiremos enseguida.

Los errores mutuos no ocurren sólo en los contratos de venta. En el caso de Raffles *vs.* Wichelhaus,[20] el contrato era para la venta de algodón a un precio fijo, y la entrega de algodón al vendedor se haría en el barco *Peerless*. Las partes no sabían que había dos barcos del mismo nombre que zarpaban para el mismo puerto en fechas diferentes (éste era el elemento "mutuo"). Dado que el precio del algodón era volátil, el valor del contrato para el comprador dependía de la fecha en la que recibiera el algodón y lo pudiera revender, lo que a su vez dependía de en cuál de los dos barcos de nombre *Peerless* se enviara el algodón. No había manera de determinar ya sea qué barco habrían elegido las partes si hubieran sabido que había más de uno o qué elección habría sido la más eficiente. Cuando ninguna de las partes es culpable, o ambas partes son igualmente culpables, de una incertidumbre incorregible en su contrato, permitir la rescisión de éste tiene sentido económico, porque en tal caso no existe el supuesto de que una parte intentaba rechazar una transacción que maximizara el valor. Sin embargo, la calificación "igualmente culpable" es importante, ya que una de las funciones de la exigibilidad del cumplimiento del contrato es la penalización de la parte que crea incertidumbre interpretativa de manera negligente.

La rescisión es el resultado habitual cuando ocurre el error mutuo, pero en ocasiones el resultado es la modificación o reformulación del contrato. Considérese el caso de Aluminum Co. of America *vs*. Essex Group, Inc.[21] En 1967 Alcoa firmó un contrato con Essex para convertir en aluminio la alúmina de Essex. Debido a que el contrato tenía un plazo largo (21 años, a elección de Essex), las partes incluyeron una cláusula de ajuste de precios basada, en parte, en el Índice de Precios al Mayoreo de los productos industriales. La electricidad es un pequeño componente del índice, pero es el insumo principal en la fabricación de aluminio. Como resultado del aumento excesivo del precio del petróleo y, por lo tanto, del costo de la electricidad (muchas plantas eléctricas funcionan con petróleo) en 1973 y 1974, el costo de cumplimiento del contrato para Alcoa aumentó mucho más rápido que el Índice de Precios al Mayoreo, provocando que se demandara la modificación. El tribunal falló a favor de Alcoa, argumentando que las partes tenían la intención de que la cláusula de ajuste de precios reflejara el aumento real de los costos de Alcoa durante la vida del contrato. Ahora bien, Alcoa había redactado la cláusula de ajuste de precios aconsejado por no menos que una figura como Alan Greenspan, un consultor económico destacado en ese tiempo. La compañía había cometido un error por descuido en el sentido económico (el cual también es

[20] 2 H. & C., 906, 159 Eng. Rep. 375 (Ex. 1864).
[21] 499 F. Supp. 53 (W.D. Pa., 1980).

jurídico) de una gran brecha entre los costos (más bajos) de evitar los errores y los costos de error (más altos). Alcoa debería haberse dado cuenta de que el Índice de Precios al Mayoreo no realizó un seguimiento de su propia estructura de costos. La aplicación de la cláusula tal cual se redactó podría haber conferido una ganancia súbita para Essex. Alcoa no se había dado cuenta de que la cláusula de ajuste tenía un vacío, y es probable que Essex tampoco se haya dado cuenta, lo que lo vuelve un caso de error mutuo. Sin embargo, hacer cumplir el contrato tal como se redactó habría creado un incentivo para una redacción más cuidadosa de las cláusulas de ajuste de precios en el futuro.

IV.5. Los contratos como seguros: las doctrinas de la imposibilidad y otras relacionadas

Hemos señalado la posible confusión que puede crearse al plantear el problema, en un caso contractual, como la decisión de hacer cumplir el contrato o no, antes que como la decisión de imponer o no una responsabilidad. Supongamos que acepto abastecer a alguien de 1 000 adminículos para el 1º de julio; mi fábrica se incendia totalmente y no puedo obtener los adminículos de nadie más a tiempo para cumplir el contrato. Supongamos además que yo no podía haber prevenido el incendio de ningún modo, así que el cumplimiento del contrato era genuinamente imposible. No se sigue de aquí que yo deba escapar a la responsabilidad por las pérdidas del comprador resultantes de mi incumplimiento. Es posible que mi compromiso haya incluido implícitamente una promesa de asegurar al comprador para el caso de que yo no pudiera entregar a tiempo los bienes prometidos.

La distinción existente entre la *prevención* y el *seguro* como métodos de minimización de la pérdida es fundamental. Una pérdida que puede evitarse mediante un gasto menor que la pérdida esperada es prevenible; no todas las pérdidas son prevenibles, pero mediante un seguro (una "indemnización" para ser más precisos, porque no es necesario involucrar a una compañía de seguros) es posible (con un costo) que se trasladen los costos creados por el riesgo de la pérdida si el riesgo se materializa. El asegurado intercambia la posibilidad de una pérdida por un costo menor, pero cierto (la prima del seguro).

Supongamos que hay una probabilidad de 1% de que ocurra un incendio que causará daños por 10 000 dólares. Como señalamos en el capítulo I, para alguien que sintiera aversión por el riesgo, este riesgo sería más costoso que una pérdida segura de su equivalente actuarial de 100 dólares. El traslado del riesgo a una compañía aseguradora eliminaría el costo adicional (la desutilidad, mas no el costo financiero) a la aversión al riesgo. Pues al reunir el riesgo de incendio de un asegurado con muchos otros riesgos de incendios con los que el asegurado no está correlacionado, la compañía de seguros transforma un riesgo en un costo (casi) seguro. A eso se debe que la compañía pueda

cobrar una prima de seguro que cubre todos sus propios costos y, sin embargo, todavía sea atractiva para el asegurado; la prima es menor que la desutilidad que éste asigna al riesgo eliminado por el seguro porque, al diversificar el riesgo, la compañía de seguros no genera esa desutilidad. Por ejemplo, la compañía de seguros podría asegurar 1 000 edificios, cada uno de ellos con una probabilidad de 1% de experimentar pérdidas de 10 000 dólares por incendio durante un periodo dado.[22] Si estas probabilidades son independientes (es decir, no relacionadas), la compañía de seguros puede estar razonablemente segura de que tendrá que pagar una suma muy cercana a 100 000 dólares por reclamaciones de los asegurados; es infinitesimal la probabilidad de que tenga que pagar el total de la pérdida asegurada: 10 millones de dólares. En cambio, la probabilidad de que el propietario del edificio individual no asegurado experimente su pérdida máxima es, como sabemos, de 1%, que no es insignificante; si siente aversión al riesgo pagará una cantidad a fin de eliminar dicho riesgo.

La compra de una póliza de seguro explícito se llama seguro de mercado, pero en algunos casos el autoaseguramiento es una alternativa mejor. Una compañía inmobiliaria es propietaria de 1 000 edificios en el ejemplo anterior; en tal caso, su "riesgo" de pérdida por incendio es una perspectiva casi segura de incurrir en un costo de 100 000 dólares y, por lo tanto, tiene poco que ganar con la compra de un seguro. En el capítulo XVI veremos que un inversionista puede reducir los riesgos asociados a un valor particular manteniendo una cartera diversificada: un conjunto de valores cuyos riesgos no se correlacionen entre sí (por lo menos parcialmente). El principio es el mismo que el del ejemplo del incendio, aunque no se le llama seguro.

Los contratos comerciales ordinarios también trasladan riesgos, y así proveen una forma de seguro. La función de traslado del riesgo o aseguramiento de los contratos se relaciona con el hecho de que un contrato (distinto del intercambio efectivamente simultáneo) compromete a una de las partes o ambas, por su propia naturaleza, a un curso de acción futuro, y el futuro es incierto. Considérese la garantía implicada de la adecuación para el uso intentado. El producto podría fallar por razones que escapan al control del vendedor, pero el vendedor podría estar en posibilidad de distribuir el riesgo de la falla a través de su base entera de clientes (cobrando un precio un poco mayor) con mayor facilidad que el consumidor, convirtiendo un riesgo en un costo seguro. En ese caso, el consumidor estará dispuesto a pagar un precio mayor a cambio de que se evite el riesgo.

O consideremos la "demora", el cargo que los transportistas hacen a los consignadores que (en el caso del transporte ferroviario, por ejemplo) retienen los carros de ferrocarril para carga o descarga por más tiempo que el establecido en su contrato con el transportista. El consignador puede escoger

[22] A fin de simplificar, omitimos las posibilidades de pérdidas intermedias.

entre dos formas del acuerdo de demora. Bajo la primera de ellas, llamada demora estricta, el consignador paga una suma fija por cada día que retenga los carros después de un periodo de gracia de dos días. Pero el cobro se condona si la demora en la devolución de los carros se debe al mal tiempo u otras circunstancias que escapen al control del consignador. Bajo el segundo acuerdo, llamado el acuerdo típico, no hay excusas por la demora, pero, por otra parte, el consignador obtiene una cantidad específica por cada carro que devuelva dentro del primer día de la llegada del carro a su muelle de carga. Así pues, la demora estricta asigna el riesgo del mal tiempo al ferrocarril, mientras que el acuerdo típico lo asigna al consignador y lo compensa por esta carga permitiéndole que le cobre al ferrocarril por la entrega temprana. La actitud del consignador hacia el riesgo, así como cualquier ventaja comparativa que pueda tener para afrontar el tiempo inesperadamente malo u otras sorpresas, determinarán la forma de la demora que escoja.[23]

Podría parecer que las partes involucradas en los contratos comprarían un seguro contra el incumplimiento y que por lo tanto no les es necesario depender de que el derecho contractual proporcione una forma de sustituto del seguro. Sin embargo, es difícil comprar un seguro contra el riesgo de ser víctima de un incumplimiento de contrato porque el riesgo y sus consecuencias no se pueden calcular con la precisión actuarial con la que lo exigen las compañías de seguros. Existe demasiada heterogeneidad entre los contratos: demasiada incertidumbre acerca de la probabilidad y las consecuencias del incumplimiento. Además, la parte asegurada contra las consecuencias del incumplimiento tendría menos incentivos para lograr que el contrato fuera exitoso. El efecto que tiene el seguro al aumentar la probabilidad de ocurrencia de un evento para la parte asegurada se llama "azar o riesgo moral".

Es más difícil comprar un seguro de responsabilidad por incumplimiento de contrato. Desde luego, el seguro de responsabilidad por herir a alguien en un accidente automovilístico por negligencia está disponible y de hecho suele ser obligatorio. Pero en ese caso el problema del azar o riesgo moral se minimiza por el hecho de que los conductores tratan de evitar accidentes por el peligro que ellos mismos corren, la poca disposición para lastimar a las personas y el temor al procesamiento penal si el accidente fue ocasionado por una infracción de tráfico grave. Sin embargo, los incumplimientos de contrato suelen ser deliberados, por lo que es difícil asegurarse contra ellos incluso al margen del azar o riesgo moral ya que su incidencia es impredecible, de modo que no se puede calcular una prima realista; un ejemplo aún más claro es el comportamiento ilícito intencional, como el que solicita una indemnización por daños punitivos a la víctima.[24]

A fin de determinar quién de las dos partes contratantes es el asegurador más barato en el caso de un percance que puede dar lugar al incumplimiento

[23] Véase Field Container Corp. *vs.* ICC, 712 F.2d, 250, 255-256 (7th Cir., 1983).

[24] Véase George L. Priest, "Insurability and Punitive Damages", 40 *Ala. L. Rev.,* 1009 (1989).

de contrato, podemos dividir los costos del aseguramiento en dos categorías: *1)* los costos de medición y *2)* los costos de transacción. Los primeros son los costos de la estimación de *a)* la probabilidad de que el riesgo se materialice y *b)* la magnitud de la pérdida si el riesgo se materializa. El producto de los dos elementos es el valor esperado de la pérdida y sirve de base para la computación de la prima de seguro apropiada que se integrará al precio del contrato. El costo de transacción principal es el de reunir el riesgo con otros riesgos a fin de reducirlo o eliminarlo; cuando el autoaseguramiento es viable, este costo podría ser menor que si debiera comprarse el seguro en el mercado.

Esta función de aseguramiento de los contratos nos ayuda a entender la doctrina de la imposibilidad y las razones relacionadas para liberar un contrato. Tal análisis explica, por ejemplo, por qué la imposibilidad física como tal no es una razón para la liberación. Si el promitente es el asegurador más barato, el hecho de que no hubiera podido prevenir el evento que le impidió el cumplimiento no debiera liberarlo.[25] A la inversa, el hecho de que el cumplimiento siga siendo físicamente posible pero sea antieconómico no debiera impedir la liberación. Si el promitente no pudo haber prevenido a un costo razonable el evento que le ha impedido cumplir su promesa, y el promisario era el asegurador más barato de la pérdida resultante, el promitente tiene un buen argumento para demostrar que no incumplió el contrato. Por lo tanto, la doctrina de la "imposibilidad" tiene un nombre errado.

La liberación se otorga rutinariamente en los contratos de servicios personales en los que la muerte del promitente impide el cumplimiento, a menos que tal individuo hubiera tenido razón para creer que su esperanza de vida era menor que lo normal para alguien de su edad (pero no hubiera prevenido al promisario). Es probable que el evento, la muerte, no sea prevenible a un costo razonable por ninguna de las partes y es poco probable que la perspectiva de la responsabilidad por el incumplimiento de contrato induzca al promitente a tomar precauciones adicionales contra la muerte. Es probable que el promisario sea el asegurador más barato porque se encuentra en mejor posición para estimar el costo si el promitente es incapaz de proveer los servicios acordados. En cambio, no se excusa al contratista que, debido a las condiciones del suelo inesperadamente difíciles, no puede completar el cumplimiento al costo que proyectó. Él debería conocer mejor que el promisario la probabilidad, y las consecuencias para los costos de la perforación, de encontrarse condiciones del subsuelo que dificultan la perforación. También debería estar en posibilidad de autoasegurarse a un bajo costo si hace muchas perforaciones en áreas diferentes y los riesgos de encontrar condiciones inesperadamente difíciles son independientes.

¿Qué sucede si un agricultor acepta, antes de la estación de la cosecha, vender su cosecha a un distribuidor de granos y la siembra es destruida por

[25] Como lo reconoció Holmes hace mucho tiempo. Véase Oliver Wendell Holmes Jr., *The Common Law,* 300 (1881).

una plaga? ¿Debiera excusarse al agricultor de toda responsabilidad? Probablemente. El agricultor tiene todos los incentivos necesarios para evitar una plaga, de modo que si cae la plaga es probable que no hubiese podido impedirlo; y el distribuidor de granos, que sin duda compra a diversos agricultores, no todos los cuales serán afectados por la plaga en la misma estación agrícola, se encuentra en una posición mejor que la del agricultor para protegerse contra el riesgo de la plaga, aunque mediante contratos de futuros ambas partes podrían ser capaces de asegurarse contra la pérdida a un costo muy bajo (véase después IV.10).

A menudo las partes incluirán en su contrato una cláusula de *force majeure* ("fuerza mayor") que especifique las circunstancias en las que se excusará el incumplimiento. Si lo hacen, las doctrinas judiciales de la imposibilidad, la impracticabilidad y otras relacionadas no deberían ser aplicables al contrato; ésas son reglas predeterminadas que las partes deberían ser libres de remplazar por sus propios términos.

Podrían multiplicarse los ejemplos de la operación de estas doctrinas,[26] pero consideremos ahora el caso relacionado en el que se impide la terminación del cumplimiento por una de las partes, de nuevo por circunstancias fuera de su control, y esa parte desea que se le excuse de un cumplimiento mayor, o aun desea que se le pague por lo que haya hecho aunque no sea lo que le imponía el contrato. Empleo a un contratista para que construya una casa y cuando la construcción está a medias se incendia el edificio. El hecho de que, sin culpa de su parte, se vea impedido para cumplir como se establecía en el contrato no debiera facultarlo automáticamente para cancelar el contrato o para que se le pague como si el edificio incendiado hubiese sido lo que yo contraté. El problema debe ser la determinación de quién de nosotros era responsable del riesgo de incendio. En ausencia de pruebas acerca de las intenciones reales de las partes, tenemos que comparar los costos relativos de la prevención o el aseguramiento contra el incendio. Al igual que un fabricante cuyos bienes son destruidos por el fuego antes de la entrega, el contratista se encuentra generalmente mejor ubicado que el propietario para la protección contra el incendio porque controla las instalaciones y es conocedor de los riesgos de incendio de los edificios en construcción. Es probable que sea también el asegurador más barato, ya que se encuentra en mejor posición que el propietario para estimar la probabilidad y las consecuencias del incendio en diversas etapas de la construcción.

[26] Véanse Richard A. Posner y Andrew M. Rosenfield, "Impossibility and Related Doctrines in Contract Law: An Economic Analysis", 6 *J. Legal Stud.,* 83 (1977); Gerhard Wagner, "In Defense of the Impossibility Defense", 27 *Loyola U. of Chi. L. J.,* 55 (1995); *cf.* Paul L. Joskow, "Commercial Impossibility, the Uranium Market, and the Westinghouse Case", 6 *J. Legal Stud.,* 119 (1977). Véase una opinión escéptica acerca de la sensatez económica de los casos decididos bajo estas doctrinas en Alan O. Sykes, "The Doctrine of Commercial Impracticability in a Second Best World", 19 *J. Legal Stud.,* 43, 73-91 (1990).

Supongamos que estalla la guerra mientras que un barco se encuentra en el puerto en ruta hacia su destino final, y la tripulación se niega a continuar el viaje a menos que el capitán del barco prometa un salario mayor. El capitán procede así, pero más tarde se niega a cumplir la promesa (como en el caso Alaska Packers') alegando que no estaba apoyada por ninguna consideración, ya que la tripulación estaba contractualmente obligada a completar el viaje. El problema consiste en saber cómo asignó el riesgo de la guerra el contrato de trabajo original. Si ese riesgo se asignó a la tripulación, entonces sus miembros habían sido compensados presumiblemente por su asunción y, por lo tanto, una promesa de pagarles salarios adicionales no estaba apoyada por ninguna consideración. Sin embargo, es probable que el capitán esté mejor capacitado para estimar la probabilidad y las consecuencias de una guerra. Por lo tanto, podemos suponer que el contrato de trabajo asignaba implícitamente el riesgo al capitán, en cuyo caso el acuerdo de la tripulación de continuar el viaje bajo condiciones de tiempos de guerra sería una consideración para la promesa de mayores salarios.

Este resultado puede parecer contrario al caso Alaska Packers', pero no lo es. En el caso Alaska Packers' no ocurrió ningún acontecimiento inesperado que excusara a los marineros de continuar trabajando por el salario convenido; su amenaza de huelga fue oportunista. Sin embargo, la guerra excusa el cumplimiento de los marineros: no la previeron en la negociación del salario, por lo que, si el capitán desea que continúen trabajando, tiene que negociar un nuevo contrato con ellos.

IV.6. Los contratos de seguros

El entendimiento de la función de aseguramiento de los contratos facilita el entendimiento del derecho contractual de seguros (formales). Veamos el principio de que los contratos de seguros deben interpretarse en contra del asegurador cuando son ambiguos. Ésta es una aplicación del principio *contra proferentem*, una regla general de la interpretación de contratos en el sentido de que en caso de duda la ambigüedad debería resolverse en contra de la parte que redactó el contrato. La doctrina se aplica con especial vehemencia a los contratos de seguros, sin embargo, y en ocasiones esto se defiende alegando que la compañía de seguros es quien puede asumir mejor el riesgo de no cobertura ocasionado por la incertidumbre interpretativa. Pero eso es incorrecto: el riesgo de la incertidumbre interpretativa no se puede cuantificar; y si una compañía de seguros no puede atribuir una probabilidad a un riesgo, no puede calcular la prima correcta que se cobrará por asumir el riesgo. Ésta es una aplicación de la diferenciación entre la incertidumbre y el riesgo calculable que se presentó anteriormente en I.1.

El principio de *contra proferentem* todavía puede ser un punto decisivo razonable (y no sólo en el caso de los seguros), argumentando que es proba-

ble que la parte que redactó el contrato se encuentre en mejor posición para evitar las ambigüedades. Pero esto no siempre es cierto. La otra parte podría tener más información acerca de la contingencia particular que originó la disputa legal: de ahí la excepción a la regla, reconocida por varios tribunales, en los casos en los que la parte que no redactó el contrato es sofisticada en términos comerciales.

Supongamos que no resulta claro a partir del contrato si el asegurado puede recuperar algo de la compañía aseguradora en los términos de la póliza, aunque obtenga el pago de daños por cuasidelitos que lo compense plenamente por el daño contra el cual se ha asegurado. Si esta ambigüedad se resuelve a favor del asegurado, ello transformará el contrato de seguro en una especie de billete de lotería que paga doble si ganas. Es probable que esto sea un aseguramiento mayor que el deseado por el asegurado cuando compró la póliza. Quizá no sea realmente un seguro en absoluto. El aseguramiento reduce la varianza del ingreso de una persona —y, por ende, el "riesgo" tal como usan este término los economistas habitualmente— porque disminuye el ingreso del asegurado en su máximo nivel por el monto de las primas y lo aumenta en su nivel mínimo (es decir, si ocurriera el desastre) por el pago estipulado por la póliza. Una póliza que sustituya con exceso al ingreso perdido aumenta, en lugar de disminuir, la varianza del ingreso, de modo que probablemente no se desea.

El principio de *contra proferentem* no siempre ha sido la regla en el caso de los seguros. En los inicios del aseguramiento formal, los contratos de seguros se interpretaban estrictamente en contra del asegurado, no contra el asegurador; cualquier cosa que hiciera el asegurado que aumentara el riesgo del asegurador tendería a verse como una "desviación" que excusaría al asegurador del cumplimiento de su parte de la negociación.[27] La tendencia de un asegurado a relajar sus esfuerzos para prevenir la ocurrencia del riesgo contra el cual se ha asegurado (azar o riesgo moral[28]) vuelve al seguro más caro de lo que sería de otro modo, pero por sí misma no es una razón en contra del seguro porque el costo añadido puede ser menor que la desutilidad del riesgo para la persona que busca un seguro. Además, no todo aumento del riesgo después del contrato se debe al azar o riesgo moral. El asegurado no tiene un control perfecto sobre las condiciones que podrían afectar ese riesgo.[29] A medida que se ampliaban los mercados de seguros, los aseguradores tenían menos necesidad de imponer el riesgo de un cambio en el riesgo al asegurado; la unión de los riesgos era suficientemente grande para permitir

[27] Véase una discusión interesante de la evolución del derecho de los seguros en Morton J. Horwitz, *The Transformation of American Law, 1780-1860*, 226-237 (1977).

[28] Véase, por ejemplo, Alma Cohen y Rajeev Dehejia, "The Effect of Automobile Insurance and Accident Liability Laws on Traffic Fatalities", 47 *J. Law & Econ.*, 357 (2004).

[29] Más adelante, en xv.2 y xv.4, discutiremos un problema paralelo en el financiamiento de las empresas.

que los aseguradores contrarrestaran el aumento del riesgo de un asegurado durante el periodo de la póliza con una disminución del riesgo de otro asegurado durante el mismo periodo. En consecuencia, la doctrina de las desviaciones se liberalizó progresivamente… lo que es una tendencia consistente con la economía cambiante del mercado de seguros.

Usted no puede asegurarlo todo: sólo algo en lo que tenga un "interés asegurable". Supongamos que A y B ven a un extraño, C, caminando por la calle, y difieren acerca de que C parezca saludable. A ofrece vender a B una póliza de seguro sobre la vida de C, y B acepta, pensando que es probable que C muera antes de lo que piensa A. Tal contrato no se hará cumplir porque B carece de un interés asegurable en la vida de C. Podríamos suponer que este resultado se debe al hecho de que la ley reprueba los contratos de apuestas, pero el requerimiento de un interés asegurable data de una época en que los contratos de apuestas eran legítimos en Inglaterra. Una explicación mejor podría ser que el contrato impone un costo externo: a C, quien no es parte del contrato.[30] Porque otorga a B un interés en que C muera lo antes posible (por supuesto, otorga a A un interés en que C viva el mayor tiempo posible, pero eso podría no consolar a C si se enterara de la existencia del contrato). Un caso real es el Da Costa *vs.* Jones,[31] una demanda para exigir que se hiciera cumplir una apuesta de que el Caballero d'Éon era en realidad una mujer; el tribunal falló que la apuesta no era exigible porque resultaba injuriosa para una tercera persona: el Caballero, quien no quería que se le considerara como mujer.

Tales casos ilustran la doctrina más amplia de que los contratos que conllevan una conducta que va contra el interés público no son exigibles, y en su mayor parte éstos son casos que imponen costos a terceros: un contrato para robar un banco es un ejemplo claro; de modo que no puede asegurarse uno mismo contra sus responsabilidades penales potenciales (véase después VII.5). A menudo se cree que los contratos de seguros imponen costos externos aun cuando el asegurado tenga un interés asegurable; por ejemplo, se cree generalmente que el seguro contra la responsabilidad externaliza (en parte) los costos de la conducción descuidada al reducir el costo de provocar una lesión para el asegurado; pero en el capítulo VI veremos que, aunque existen pruebas de que los conductores manejan con menos precaución cuando están asegurados porque se sienten más seguros,[32] no se sigue que éstos externalicen los costos de los accidentes adicionales.

Las compañías de seguros toman diversas medidas para protegerse contra el azar o riesgo moral. Una de ellas es negarse a asegurar la propiedad por más de lo que vale, o, como se mencionó en la sección anterior de este capí-

[30] Samuel A. Rea Jr., "The Economics of Insurance Law", 13 *Int'l Rev. Law & Econ.,* 145 (1993).

[31] 2 Cowp. 729, 98 Eng. Rep. 1331 (K. B., 1778).

[32] Cohen y Dehejia, véase la nota 28.

tulo, asegurarse contra la incidencia de una conducta que no se puede predecir. Otra medida es excluir la cobertura de las pérdidas que el asegurado podría prevenir muy fácilmente. Por ejemplo, la "Póliza del Bloqueo de los Joyeros", una póliza de seguro estándar que protege a los joyeros contra una amplia gama de riesgos de pérdida o daño, incluye una excepción para los casos en los que las joyas se pierden "mientras estaban dentro de o sobre cualquier vehículo", a menos que "en el momento en que ocurra la pérdida realmente esté dentro de o sobre dicho vehículo […] un empleado fijo del Asegurado" y "la propiedad asegurada esté bajo la custodia personal estricta y bajo el control directo de [el empleado]". El riesgo de pérdida por robo de las joyas que se dejan sin supervisión en un vehículo es tanto muy alto como fácil de reducir al exigir que no se dejen sin supervisión.[33]

El azar o riesgo moral en los seguros no se limita al comportamiento de los asegurados. Consideremos los incentivos de un asegurador de responsabilidades cuando se promueve un juicio contra su asegurado que tiene un valor esperado mayor que el de los límites de la póliza. La compañía de seguros sabe que tendrá que pagar sólo lo establecido. En lugar de pelear el juicio, por ello, el asegurador (que en los términos habituales de la póliza paga los gastos del litigio y por ello se reservará de ordinario el derecho de manejar la defensa) deseará arreglarlo aun si esto exige convenir en pagar por más que el límite de la póliza (ya que la diferencia será soportada por el asegurado, y el asegurador evitará el gasto del juicio), e incluso si el litigio produjera un resultado mejor para el asegurado. A fin de impedir que la compañía aseguradora abandone al asegurado de esta manera, la ley interpreta que la póliza impone al asegurador una obligación fiduciaria a favor del asegurado; es decir, se exige que el asegurador trate al asegurado como se trataría a sí mismo.

Un problema claro del aseguramiento es el de la selección adversa. Si un asegurador ofrece asegurar contra algún riesgo a todos los que lleguen, a un precio uniforme, será mucho más probable que acepten la oferta las personas con riesgos superiores al promedio que las personas con riesgos menores al promedio, porque el efecto del precio uniforme será que estas últimas subsidien a las que están por encima del promedio, dando como resultado un equilibrio de la reunión (véase I.4). Si quien se encuentra por debajo del promedio rechaza la oferta, el asegurador deberá aumentar el precio del seguro, y el efecto será que los clientes que se encuentren por debajo del promedio del grupo que antes estaba por encima del promedio estarán subsidiando a quienes se encuentren por encima del promedio, de modo que también se saldrán.

Los aseguradores pueden tratar de impedir este ciclo descendente cobrando precios diferentes a los asegurados contra riesgos diferentes. Pero esto no funcionará si la información acerca del riesgo relativo es exclusiva de los asegurados. Una alternativa popular consiste en impedir que se salgan

[33] Véase A. M. I. Diamonds Co. *vs.* Hanover Ins. Co., 397 F.3d, 528 (7th Cir., 2005).

las personas de riesgo por debajo del promedio, como ocurre con el seguro médico basado en el empleo, que cubre a todos los empleados de la empresa que compra la póliza de seguros independientemente de la salud de cualquier empleado en particular. Una alternativa controvertida consiste en excluir a clases de riesgo enteras, como las del embarazo y el sida (en el caso del seguro médico). Estas exclusiones son criticadas por discriminatorias (en el caso del seguro médico proveído por el empleador, una exclusión por embarazo se considera como una discriminación sexual); sin embargo, no se critica la exclusión de la muerte por suicidio de la cobertura de una póliza de seguro de vida, lo que tiene sus bases: sin la exclusión, las personas que planean suicidarse podrían obtener una ganancia súbita para sus beneficiarios a costa de los demás miembros del grupo de seguros a través de la compra de un seguro de vida en la víspera de su suicidio (evitando así el costo significativo de la prima).

Imaginemos que los avances en las pruebas genéticas permiten que las compañías de seguros de vida y de salud identifiquen a los asegurados de alto riesgo y que se les niegue el seguro o se les cobre una prima más alta en comparación con otros asegurados. En el límite, los mercados de seguros colapsarían porque no habría incertidumbre contra la cual asegurarse. Esto sería un resultado eficiente, ya que eliminaría los costos administrativos no poco considerables que impone el mercado de seguros (incluyendo a los litigios), pero agravaría las desigualdades de ingresos y de riqueza.

Los seguros son una industria regulada (en los Estados Unidos la regulan los comisionados estatales de seguros). La razón principal es asegurar la solvencia de las compañías de seguros. Esto es una preocupación porque mientras que en la mayoría de las industrias los vendedores incurren en los costos antes de recibir sus ingresos, en el caso de los seguros sucede lo contrario, lo que puede crear incentivos perversos, en especial cuando las pérdidas contra las que se aseguran son eventos con una probabilidad de ocurrencia muy baja. Imaginemos un riesgo que sólo tiene una probabilidad anual de materialización de 1%. La compañía de seguros podría cobrar las primas a sus asegurados durante muchos años antes de tener que pagar alguna reclamación, por lo que podría considerar conveniente el pagar dividendos muy altos a sus accionistas durante este periodo, incluso si eso significara que no pudiera pagar las reclamaciones de sus asegurados en su totalidad. Los reguladores de seguros exigen que las compañías de seguros mantengan un nivel de efectivo o de reservas equivalentes de efectivo que aseguren la capacidad de las empresas para pagar las reclamaciones que se presenten.

En el colapso financiero de 2008 ocurrió una falla regulatoria dramática. American International Group, que probablemente era la compañía de seguros más grande del mundo en ese entonces, había emitido decenas de miles de millones de dólares en "permutas de incumplimiento crediticio", una forma de seguro de crédito que le asegura al tenedor de bonos que en caso de que

exista un hecho desencadenante, como lo es una pérdida específica en el valor de mercado del bono, el emisor de la permuta resarcirá la pérdida de los tenedores de bonos. El precio de las permutas fue muy bajo, porque prácticamente nadie pensaba que fuera remotamente probable que el mercado de bonos sufriría una pérdida en valor lo suficientemente grande como para permitir que se hicieran reclamaciones considerables contra AIG, el emisor de las permutas. Pero gran parte de los bonos asegurados resultaron ser valores respaldados por hipotecas que experimentaron pérdidas de valor dramáticas en 2008 (véase después XVI.14) y, en consecuencia, AIG se encontró con pasivos demoledores ante los compradores de las permutas. Debido a que los comisionados estatales de seguros no regulaban las permutas de cobertura por incumplimiento crediticio como contratos de seguros, los compradores de permutas habían tenido que hacer sus propios arreglos para asegurar la solvencia de los emisores. Por lo general, eso significaba que se exigiría al emisor que depositara una garantía. Sin embargo, debido a la calificación crediticia destacada de AIG, la mayoría de sus clientes de permutas no insistieron en la garantía, y dado que no esperaba la caída del mercado de bonos, AIG no había acumulado reservas contra tal eventualidad. La incapacidad de AIG para cumplir con sus obligaciones de permutas puso en riesgo la solvencia de sus clientes, entre los cuales se encontraban empresas financieras importantes como Goldman Sachs. El Departamento del Tesoro intervino mediante la indemnización de las obligaciones de AIG y haciéndose cargo de la empresa.[34]

IV.7. El fraude

Aun cuando desde la firma del contrato no haya ocurrido nada que vuelva antieconómico el cumplimiento, podría permitirse la liberación sin responsabilidad cuando se refute la presunción de que el cumplimiento generaría un intercambio que aumente el valor, como ocurre cuando se demuestra que el promisario indujo la promesa mediante una mentira. Pero supongamos que el promisario no miente, sino que sólo dejó de revelar la información que tiene y que habría destruido el trato si la hubiera comunicado al promitente. En Laidlaw *vs*. Organ,[35] un comerciante de Nueva Orleans llamado Organ, enterado del Tratado de Gante (que terminó la guerra de 1812) antes de que la noticia se hiciera pública, ordenó una gran cantidad de tabaco a la empresa Laidlaw al precio de mercado de entonces. Cuando el público se enteró del tratado, que terminaba con el bloqueo naval británico de Nueva Orleans, el pre-

[34] En ocasiones se concede el pago de daños punitivos por el incumplimiento de los contratos de seguros, una excepción a la regla general de que no se concede el pago de daños punitivos en los casos de contratos. Véase IV.12 después.

[35] 15 U.S. (2 Wheat.) 178 (1817), discutido en Anthony T. Kronman, "Mistake, Disclosure, Information, and the Law of Contracts", 7 *J. Legal Stud.*, 1, 9-18 (1978).

cio del tabaco aumentó de 30 a 50%, y Laidlaw trató de retractarse. La Corte Suprema sostuvo que no podía hacerlo. Si Organ, cuando colocó su orden, le hubiese dicho a Laidlaw, como podría haberlo hecho fácilmente, que el tratado había sido firmado, Laidlaw habría insistido en un precio mayor, y puesto que no le habría costado nada a Organ la comunicación de la información que ya tenía, podría pensarse que el fallo de la Corte Suprema le otorgaba simplemente una ganancia indecente. Pero esto no es correcto. Aunque la información, una vez obtenida, podría transmitirse a bajo costo a otra persona, a menudo no es barata su obtención, y si no permitimos que las personas se beneficien de la información manteniéndola en secreto, no tendrán incentivos para obtenerla en primer lugar, y la sociedad será la perdedora. (*Cf.* la discusión de las patentes y las marcas registradas en el capítulo XI.) Por supuesto, éste es un argumento muy pobre si la información se mantiene secreta para siempre, pero Organ no hizo eso. Al colocar su orden con Laidlaw, aumentando así la demanda de tabaco, hizo que el precio aumentara, aunque no tanto como ocurrió cuando la información se reveló por completo. El proceso del ajuste de los precios a la información de que la guerra había terminado se inició antes a causa de la transacción de Organ.

La mentira es diferente. El mentiroso hace una inversión positiva en la fabricación y diseminación de la información falsa. Esta inversión es un desperdicio desde un punto de vista social, de modo que naturalmente no lo recompensamos por su mentira. Éste es un caso intermedio: el vendedor de una casa, A, sabiendo que ésta se halla infestada de termitas, no se lo dice a su comprador B. Podría argüirse (la autoridad judicial está dividida sobre este punto) que A tiene una obligación de informar o, en el lenguaje legal, que el hecho de que no lo haga constituye una omisión que amerita acción. Es probable que A no haya invertido en el descubrimiento de que la casa tiene termitas. Adquirió la información de manera gratuita, como un subproducto del hecho de vivir allí. Además, esta información beneficia a menos personas que la información acerca del valor del tabaco: hay menos compradores potenciales para una casa que para todo el mercado del tabaco. Por lo tanto, el beneficio de la información es menor y la necesidad de proveer una protección legal a fin de inducir su producción es en consecuencia también menor.

Además, a menos que haya una obligación de decir la verdad en el caso de las termitas, el comprador invertirá en una inspección de termitas o negociará la inclusión, en el contrato, de una cláusula que garantice que la casa está libre de ellas. Estos costos se evitan mediante la imposición al vendedor de la obligación de revelar la información que él obtuvo sin costo alguno.

Con el ejemplo de las termitas pasamos al campo del fraude al consumidor, un problema más grave que el del fraude en los contratos comerciales a causa de la disparidad de los recursos y los conocimientos entre el vendedor y el comprador en el contexto de los productos de consumo; de las deficiencias

cognitivas que, como han descubierto los psicólogos, juegan un papel importante en el comportamiento del consumidor; y de la dificultad de diseñar remedios legales eficaces cuando los intereses en juego son pequeños, como suelen ser en muchas disputas entre vendedores y consumidores. Los problemas psicológicos y de remedios se discuten con mayor detalle en los capítulos XIV y XXII, respectivamente, aunque advierto aquí que el problema remedial podría no ser difícil cuando el cumplimiento del consumidor sigue al del vendedor, como ocurre en el contrato típico en abonos. En tal caso, el consumidor defraudado no tiene que iniciar un juicio legal. Puede rescindir. Por supuesto, la eficacia de este remedio depende del momento en que el consumidor descubra el fraude y también de la aplicabilidad de la doctrina del tenedor en el momento oportuno, que se discute en la sección siguiente.

Podría pensarse que la disponibilidad de remedios legales carecería de importancia; porque si una empresa está ganando ventas a sus rivales mediante declaraciones falsas acerca de su producto (o acerca de los productos de sus rivales), los rivales tienen un incentivo para exponer las declaraciones falsas ante el consumidor. Aún hay más: las asociaciones comerciales establecen estándares de calidad en los que los consumidores pueden confiar; las marcas comerciales tienen una función similar (véase después XI.7); los medios informan a los consumidores acerca de los méritos de bienes particulares; y la tienda de departamentos, como un comprador experto de los bienes de muchos fabricantes rivales, ayuda al consumidor a escoger sensatamente entre productos rivales.

El proceso por el cual corrigen los competidores las impresiones engañosas creadas por uno de sus miembros no opera, con todo, instantáneamente, y en el ínterin las ganancias obtenibles por el vendedor fraudulento podrían superar a largo plazo a cualesquiera costos derivados de la pérdida de reputación, sobre todo si el vendedor puede abandonar el mercado a bajo costo. Además, si un vendedor fraudulento está desviando sólo una pequeña cantidad de operaciones de cada uno de sus rivales, ninguno de ellos tendrá un incentivo para tomar medidas costosas a fin de corregir sus declaraciones falsas, aunque el total de ventas desviadas por el vendedor fraudulento podría ser muy grande. Y existe un problema más grave, conocido como "la maldición de la corrección del sesgo".[36] Si un vendedor expone la publicidad engañosa de sus competidores, éstos corregirán su publicidad a fin de evitar que las ventas se desvíen hacia él. Pero, una vez que lo hagan, el vendedor no obtendrá beneficio alguno por haberlos expuesto. De hecho, su situación podría empeorar, junto con la de todos sus competidores, debido a que es probable que la publicidad falsa haya elevado las ventas.

La asociación comercial es sólo una respuesta parcial. Los miembros de una industria no tienen a menudo fuertes incentivos para apoyar la actividad

[36] Véase, por ejemplo, Ann-Sophie Vandenberghe, "Behavioral Approaches to Contract Law", en *Contract Law and Economics*, 401, 417 (Gerrit De Geest [comp.], 2011).

de la asociación comercial, porque el vendedor que no contribuye nada a la campaña de la asociación contra el fraude obtendrá no obstante de la campaña los mismos beneficios que los otros vendedores, y sin costo alguno, siempre y cuando o haya tantos vendedores que la contribución a la campaña de uno de ellos no sea esencial o que, aunque no haya muchos vendedores, el vendedor en cuestión sea muy pequeño en relación con los demás. Éste es el conocido problema del gorrón.

Además, no todas las industrias o mercados son competitivos. Un monopolista o un cártel podrían tener un incentivo mayor que el de una empresa en una industria competitiva para declarar falsamente las calidades de su producto. El efecto del fraude del monopolista para desviar las ventas hacia él se dispersaría entre gran número de bienes sustitutos, ninguno de los cuales se verá grandemente afectado;[37] por lo tanto, es menos probable que en el caso competitivo cualquier vendedor tenga un incentivo para combatir la declaración falsa. Un argumento relacionado, pero más general, es que cuando un hecho embarazoso acerca de un producto se aplica igualmente a todas las marcas del producto, ninguno de los productores tendría un fuerte incentivo para revelarlo, aunque la industria sea competitiva. Si un fabricante de cigarros anuncia que fumar es bueno para la salud, otros fabricantes de cigarros no tendrán ningún incentivo para refutar esa declaración. Y, dado que no hay sustitutos cercanos para el tabaco cuyos productores pudieran prever un gran aumento de sus ventas si persuaden a los consumidores de que no les conviene fumar, ningún otro fabricante tendrá tampoco un fuerte incentivo para refutar la declaración.

De los problemas que se expresan en esta sección no se concluye que debiera haber una obligación *general* de los vendedores de bienes de consumo de revelar a los consumidores la información importante. Si la característica relevante del producto es tal que el comprador la puede determinar mediante la inspección o el manejo casuales en el momento de la compra —la suavidad de un suéter de cachemira, por ejemplo—, entonces sería redundante requerir que el vendedor la revelara. A menudo, la determinación de las características de un producto requiere el uso efectivo antes que la mera inspección o manejo antes de la venta (el efecto blanqueador de un líquido casero es un ejemplo). Aun así, si el producto es barato y se compra repetidamente, el costo de la determinación de la característica seguirá siendo muy bajo para el consumidor: el costo de la primera compra. A veces, *sólo* el consumidor tiene la información requerida, porque el desempeño del producto puede depender de los gustos del consumidor, los que pueden ser desconocidos por el fabricante. Sólo el consumidor sabe si el suéter es suficientemente suave o el melón suficientemente maduro para él o ella.

[37] Por definición, no hay ningún sustituto *cercano* para el bien del monopolista; de otro modo, no se trataría de un caso de monopolio.

Pero, evidentemente, a menudo el consumidor no puede determinar una característica del producto a un costo bajo. Vemos algunos ejemplos cuando el producto se compra infrecuentemente y la característica no puede descubrirse mediante la inspección o el manejo en el momento de la compra; o bien, aunque el producto se compre con frecuencia, es muy caro, como un automóvil; o la característica no puede descubrirse ni siquiera mediante el uso repetido o prolongado, como ocurre con la calidad de los consejos de un libro acerca de cómo ganar un millón de pesos en el negocio inmobiliario. Sin embargo, en estos casos (que ejemplifican lo que los economistas llaman "bienes de credibilidad") las presiones competitivas hacen a menudo que los vendedores ofrezcan garantías (exigibles) de características particulares de sus productos. Una garantía no es una revelación. Es algo mejor: una garantía de resultados, lo que hace innecesaria la revelación de información. Un fabricante de televisores que garantiza que el tubo durará tres años no necesita revelar cuál es la vida útil del tubo. Si se descompone después de sólo un año, el consumidor no verá empeorar su situación, ya que obtendrá un repuesto sin cargo alguno.

Si es claro que las partes intentan que el vendedor asuma el riesgo de la ignorancia del consumidor acerca de un atributo particular del producto, no habrá necesidad ni siquiera de una garantía explícita: el derecho contractual desempeñará su conocida función de economizar en los costos de transacción interpolando una garantía en el contrato de venta. Una garantía implícita de que una lata de sardinas se encuentra en buen estado es una alternativa barata para una garantía expresa de adecuación para el consumo humano y para un requerimiento legal de que el vendedor revele las características sanitarias del producto.

Si el vendedor miente acerca de su producto, antes que dejar simplemente de revelar información negativa al respecto, su conducta es ilegal, aunque el comprador pudiera descubrir la mentira a un costo muy bajo. Esto tiene sentido económico. A ofrece vender a B una caja de dulces. B le pregunta a A si es necesario abrir la caja para ver si contiene dulces, y A contesta que no es necesario, que B puede confiar en su palabra. Así que B compra la caja sin inspeccionarla, y cuando llega a su casa descubre que contiene comida para canarios en lugar de dulces. Si la mentira no estuviera sujeta a acciones, según la teoría de que B podría haber evitado fácilmente las consecuencias mediante una inspección, entonces los B de este mundo tendrán que inspeccionar y los costos agregados de la inspección podrían ser sustanciales. En cambio, los costos para A de no mentir serían negativos (compárese después VI.15).

Sin embargo, si los consumidores prestan poca atención a las revelaciones de los vendedores, si las revelaciones se redactan con astucia para evitar proporcionar advertencias inteligibles sin que se lleven a juicio las acusaciones de fraude, si la competencia no genera información adecuada para los

consumidores porque los vendedores competidores se benefician con la ignorancia de los consumidores, en resumen, si los consumidores suelen ser como corderos que se llevan al matadero, el derecho contractual no podrá optimizar el funcionamiento de los mercados de bienes de consumo. Como veremos en los capítulos siguientes, podría requerirse de otra norma jurídica.[38]

IV.8. El derecho de la agencia

Un agente es una persona que actúa en nombre de otra, el principal. La relación de la agencia es contractual y el derecho de la agencia es propiamente visto como parte del derecho contractual. En el tipo de relación que el derecho llama fiduciaria o confidencial, la obligación de revelar es mucho mayor que en una relación comercial ordinaria y la mayoría de los agentes (abogados, contadores, corredores, administradores, etc.) son agentes fiduciarios. Éste es un principio importante del derecho de la agencia. Al agente se le paga para que trate al principal como se trataría a sí mismo; para que sea su otro yo. El principio fiduciario es la respuesta de la ley al problema de los costos desiguales de la información. Ese principio le permite a usted contratar a alguien que tiene mejor información para tratar en su nombre con otros que tienen mejor información. Al imponer una obligación de máxima buena fe y no la obligación contractual estándar de la buena fe ordinaria, el principio en cuestión minimiza los costos de la autoprotección para el principal del fiduciario. Esto es especialmente importante en los contextos en que el principal es totalmente impotente para protegerse por sí mismo: podría ser un niño, por ejemplo, incluso un nonato. (Más adelante, en XV.9, se discuten las obligaciones fiduciarias dentro de una empresa.)

Pero el derecho de la agencia se preocupa menos por las obligaciones del agente hacia su principal (y viceversa) que por la obligación del principal ante los terceros que tratan con su agente. Bajo la doctrina de que el superior es responsable, el principal puede ser considerado responsable del cuasidelito de un agente aunque no tenga ninguna culpa (véase después VI.8). Y está generalmente vinculado por los contratos que su agente celebre, siempre que el agente actúe dentro del campo de su autoridad. Pero hay algunas excepciones. Por ejemplo, si la persona con la que el agente celebra el contrato sabe que éste está ebrio, el principal no quedará vinculado si el agente no estaba efectivamente autorizado para celebrar el contrato. En cambio, bajo el principio de la autoridad aparente, el principal podría quedar vinculado por un contrato negociado por su agente, aunque éste actúe fuera del campo de su autoridad, siempre que parezca estar autorizado. En efecto, si alguien que no

[38] Véase después VI.6 (la responsabilidad por daños ocasionados por productos defectuosos o desproporcionadamente peligrosos), y más adelante XIV.2 (el fraude al consumidor).

es un agente en absoluto se hace aparecer como si lo fuera, y la persona a quien el agente ha representado como su principal sabe de la falsa representación pero no previene a la persona con la que el "agente" celebra el contrato, el "principal" quedará vinculado. En todos estos casos, la ley asigna una responsabilidad a la persona que puede evitar el error al menor costo (por ejemplo, el principal en el caso de la autoridad aparente porque tiene sobre el agente un control mayor que el que tiene el tercero).

Pero supongamos ahora que el tercero celebra un contrato con un agente ignorando que lo es; ¿quedará vinculado el principal (a quien la ley llama un "principal no revelado") y, por lo tanto, podrá ser demandado por el tercero por incumplimiento de contrato a pesar de que el tercero ignore su existencia? Sí, porque es probable que el principal pueda prevenir el incumplimiento (o al menos reducir su probabilidad). Y, dado que el agente se declarará insolvente con frecuencia, si se exime al principal se le permitiría externalizar el costo del incumplimiento hacia la otra parte que celebró el contrato.[39]

IV.9. Coacción, contratos impresos, poder de negociación e inconsciencia

Como se señaló anteriormente, la coacción es una defensa contra una acción de incumplimiento de contrato. En su sentido original, la coacción implica una amenaza de violencia. A le apunta a B con una pistola, diciéndole: "El dinero o la vida"; B acepta la primera parte de esta oferta entregando su dinero. Pero un tribunal no hará cumplir el contrato resultante. La razón no es el hecho de que B no esté actuando por su propia y libre voluntad. Por el contrario, B estaba sin duda en extremo interesado en aceptar la oferta de A. Sin embargo, el cumplimiento forzoso de tales ofertas disminuiría el producto social neto al canalizar recursos hacia la formulación de amenazas y los esfuerzos por protegerse contra ellas.

Consideremos este ejemplo: Un barco se incapacita y la tripulación huye, dejando a bordo sólo al patrón, quien vamos a suponer que es también el representante del dueño del barco autorizado para tratar con cualesquiera rescatadores que pasen casualmente por el lugar. Un remolcador operado por una compañía de salvamentos se aproxima al barco y el capitán del remolcador le ofrece al patrón del barco un contrato para rescatar éste a un precio igual a 99% del valor del barco y de su carga. Si el patrón del barco firma el contrato, ¿debiera el dueño del barco respetar el contrato? El derecho marítimo dice que no. Se trata de una situación de monopolio bilateral, con la complicación adicional de que el escenario de emergencia puede hacer que la negociación sea inviable y, así, los costos de transacción sean prohibitivos. Este problema

[39] Sobre esta cuestión y esta sección, en general, véase Eric Rasmusen, "Agency Law and Contract Formation", 6 *Am. Law & Econ. Rev.*, 369 (2004).

puede evitarse mediante una regla —la piedra miliar de las reglas de salvamento marítimo— que establece que el rescatador tiene derecho a un pago razonable por el salvamento del barco, pero que un contrato celebrado después de que el barco se mete en un atolladero será sólo una prueba de lo que es ese precio razonable.[40]

Mediante una moderada extensión del significado, la coacción puede emplearse también para describir el uso de una amenaza de incumplimiento para inducir una modificación de los términos del contrato en casos como Alaska Packers' Assn. *vs*. Domenico, en los que el promisario carece de remedios legales adecuados. Las operaciones de rescate ofrecen ejemplos adicionales. A descubre a B perdido en medio de una tormenta de nieve y se niega a ayudarlo a menos que B prometa darle a A toda su riqueza. Es posible que también aquí debiera excusarse a B del cumplimiento de su promesa; si permitimos las ganancias monopólicas en las operaciones de rescate, podría atraerse al negocio de los rescates una cantidad excesiva de recursos. Adviértase que ambos casos, el de la tormenta de nieve y el de Alaska Packers, son de monopolio temporal "situacional".

Cuando la transacción ocurre entre una gran corporación y un individuo ordinario, nos vemos tentados a invocar la analogía de la coacción y comparar al individuo con el sujeto inerme que se ve obligado a firmar un pagaré con una navaja en la garganta —especialmente si su contrato con la corporación es un contrato estándar impreso o el consumidor es una persona pobre—, y concluir que los términos del contrato son coercitivos. Muchos contratos (los contratos de seguros son un buen ejemplo) se ofrecen sobre la base de tómalo o déjalo. El vendedor le entrega al comprador un contrato estándar impreso que establece, a veces con un detallismo abrumador, las obligaciones respectivas de las partes. El comprador puede firmarlo o no, según desee, pero no hay negociación sobre los términos. Podemos pasar fácilmente de la observación de que no hay negociación a la conclusión de que el comprador carecía de una elección libre y, por lo tanto, no debiera quedar vinculado. Pero hay una explicación inocente para estos "contratos de adhesión": que el vendedor está tratando de evitar los costos de la negociación y la redacción de un acuerdo separado con cada comprador. Estos costos, el mayor de los cuales es probablemente el costo de supervisar a los empleados y agentes que realizan las negociaciones contractuales efectivas a nombre de la compañía, tienden a ser elevados para una compañía grande que tiene muchos contratos. Es difícil imaginar una economía de consumo moderna que funcione sin contratos impresos.

Consistentes con la explicación inocente, los compradores grandes y sofisticados, así como los consumidores individuales, realizan a menudo com-

[40] Véanse William M. Landes y Richard A. Posner, "Salvors, Finders, Good Samaritans, and Other Rescuers: An Economic Study of Law and Altruism", 7 *J. Legal Stud.*, 83, 100-105 (1978); Paul Hallwood y Thomas J. Miceli, "Murky Waters: The Law and Economics of Salvaging Historic Shipwrecks", 35 *J. Legal Stud.*, 285 (2006).

pras mediante contratos impresos. En ocasiones, hacer esto conduce a lo que los abogados contractuales llaman "la batalla de las formas". A podría firmar, y enviar a B, un contrato impreso que constituye una oferta, y B, en lugar de firmar la forma de A, podría firmar, y enviar a A, su propia forma impresa. Bajo la regla de la "imagen de espejo" del derecho común, si la forma de B no es idéntica a la de A, no sería una aceptación de la oferta de A, sino que se trataría como una contraoferta. El Código Comercial Uniforme (CCU) adopta una postura más liberal. Si A, al recibir la forma de B, empieza a cumplir, esto se toma como una aceptación de la oferta de B, aunque contenga términos muy discrepantes, porque A podría haberle informado a B que tal oferta era inaceptable en lugar de empezar a cumplir. De este modo, el enfoque del CCU reduce los costos de transacción: al tomar como aceptación el comienzo del cumplimiento, se elimina la necesidad de una comunicación.

La explicación siniestra del contrato estándar impreso es que el vendedor se niega a tratar por separado con cada comprador porque el comprador no tiene más remedio que aceptar sus términos. Si, empero, un vendedor ofrece términos poco atractivos, ¿no ofrecerá términos más atractivos un vendedor rival, deseando las ventas para él mismo, y continuará el proceso hasta que los términos sean óptimos? Todas las empresas de la industria podrían considerar económico el empleo de contratos estándares impresos y negarse a negociar con los compradores. Pero aquí lo importante no es si hay regateo en cada transacción, sino si la competencia obliga a los vendedores a incorporar en sus contratos estándares algunos términos que protejan a los compradores. Que lo haga o no podrá depender, no obstante, de la capacidad de los consumidores, como se sugirió anteriormente; y más aún de los incentivos de los consumidores, como se sugiere más adelante.

En el monopolio, por definición, el comprador no tiene buenas alternativas para tratar con el vendedor, quien se encuentra, por lo tanto, en situación de obligar al comprador a aceptar términos que en un mercado competitivo serían mejorados por otro vendedor. No se concluye de aquí que el comprador será indiferente a los términos del contrato ofrecido por el vendedor. Por el contrario: dado que un producto monopolizado tendrá un precio mayor que bajo la competencia, los compradores potenciales invertirán más tiempo y esfuerzos, no menos, en la búsqueda, y una forma de la búsqueda del consumidor es la lectura cuidadosa de los términos de un contrato. Le conviene al consumidor leer el contrato incluso si sabe que el vendedor monopólico no negociará (regateará) con él, ya que debe decidir entre comprar el producto o un sustituto. El hecho de que un producto esté monopolizado no lo convierte en un bien de primera necesidad.

Una vez dicho todo esto, los contratos estándares impresos que se utilizan en las transacciones de los consumidores sí tienden a ser unilaterales en contra del consumidor; evidentemente, no se puede confiar en que la competencia produzca el estándar óptimo. Dado que las trampas ocultas en el lenguaje de

un contrato sólo surgen en las raras ocasiones en las que hay una disputa legal, el beneficio derivado de un estándar "bueno" para el consumidor es mínimo y, por lo tanto, es poco probable que influya de manera significativa en su decisión de comprar el producto del vendedor. En cambio, el vendedor, al conocer mucho mejor la probabilidad y las consecuencias de dicha disputa, anticipará una pequeña ganancia de imponer a sus clientes un estándar "malo". No se puede confiar en que la competencia elimine esta asimetría porque los beneficios del estándar "bueno" para el consumidor son demasiado reducidos para superar los costos de la información de hacer que esos beneficios sean un factor de venta real. Además, el consumidor potencial se deprime cuando se le recuerda la posibilidad de un juicio. ("No se preocupe: si lo demando, el contrato lo protegerá.")

Sin embargo, después de todo, ¡el contrato estándar impreso unilateral puede ser óptimo![41] Es más probable que se disuada a un vendedor de comportarse de manera oportunista por consideraciones de reputación que a un consumidor. El consumidor no tiene una reputación que perder a menos que su comportamiento oportunista en una transacción específica se dé a conocer en todo el mercado. Una forma de restablecer el equilibrio es sesgar los términos del contrato a favor del vendedor. Un contrato unilateral no implica que la transacción sea necesariamente unilateral, sino sólo que el vendedor será discreto con respecto a la forma de tratar al consumidor. Se puede esperar que un vendedor al que le preocupa su reputación trate mejor a los consumidores de lo que exige el contrato, pero su derecho a depender del contrato tal como está escrito lo protegerá de los compradores oportunistas. Por tanto, las partes informadas *ex ante* podrían preferir un contrato unilateral como un mecanismo más barato para inducir resultados eficientes, si surgieran contingencias durante el cumplimiento del contrato, antes que incurrir en los costos de un contrato más elaborado que tratara de especificar y prohibir cualquier forma de comportamiento oportunista en el que el comprador se pudiera involucrar.

Los términos contractuales que parecen desventajosos para los infortunados consumidores pueden beneficiarlos aún más al reducir los costos de los vendedores. (Como veremos en el capítulo IX, incluso un monopolista reducirá su precio si sus costos disminuyen.) Tenemos un ejemplo en la venta a crédito cuando el comprador acepta que el vendedor podrá descontar el pagaré del comprador con una compañía financiera. Según el derecho común, la compañía financiera, como tenedora de buena fe (o en el debido curso), podría cobrarlo libre de toda defensa que el comprador pudiera haber hecho valer en un juicio de cobro promovido por el vendedor. Por lo tanto, si compró un escritorio en una mueblería y el escritorio resultó defectuoso, pero la

[41] Véase Lucian A. Bebchuk y Richard A. Posner, "One-Sided Contracts in Competitive Consumer Markets", 104 *Mich. L. Rev.*, 827 (2006). También véase *Boilerplate: The Foundation of Market Contracts* (Omri Ben-Shahar [comp.], 2007).

tienda ha descontado el pagaré con una compañía financiera, usted tendría que pagar la cantidad total del pagaré y se quedaría sin un derecho para demandar a la tienda por incumplimiento de la garantía. Tal disposición disminuye el costo del financiamiento para los compradores en abonos al volver más baratos y más seguros los juicios de cobro.[42] Si no existiera esa disposición, este costo (un costo soportado por el consumidor, por lo menos en su mayor parte; véase antes en III.13) sería mayor. No es obviamente mejor para el consumidor decidir pagar más por un producto que su decisión de renunciar a uno de sus remedios legales contra el vendedor. También adviértase que se favorece a los consumidores prudentes, aquellos cuyo incumplimiento es poco probable, porque la pérdida de la defensa contra el tenedor de buena fe (o en el debido curso) no significa un costo para alguien que no es demandado.

Supongamos que un contrato en abonos establece que una falta de pago facultará al vendedor para recuperar el bien, por pequeño que sea el saldo no pagado del pagaré por el comprador, y quedarse con el monto total de la reventa del bien a alguien más.[43] Si el impago ocurre hacia el final del plazo del pagaré, recuperar el bien conferirá una ganancia súbita al vendedor, quien casi ha recibido el precio total del bien, incluidos los intereses. Pero si el impago ocurre al principio, el vendedor experimentará una pérdida inesperada; sólo ha recibido una pequeña parte del precio, demasiado poco para cubrir la depreciación del bien y los costos de la recuperación. (Esto supone, pero realistamente, que el vendedor no será capaz de cobrar directamente al comprador el saldo no pagado demandándolo.) La limitación de las ganancias inesperadas provenientes de los impagos finales obligaría a los vendedores a exigir enganches mayores o pagos de abonos iniciales mayores o a cobrar precios más altos, a fin de protegerse contra las pérdidas aleatorias derivadas de los impagos tempranos. Los consumidores (en particular los prudentes) que no puedan dar grandes enganches o grandes pagos de abonos iniciales se verían perjudicados por el cambio de la forma contractual.

Por consiguiente, el caso es totalmente diferente del de "el dinero o la vida". Este último representa una clase de transacciones que empeora la situación de los "compradores" *ex ante*; en cambio, los consumidores provistos de

[42] Supongamos que, en ausencia de tales disposiciones, las compañías financieras impusieran a los vendedores promesas de indemnización por cualesquiera pérdidas que sufran como resultado de que ellas no puedan cobrar su pagaré porque el comprador tiene una buena defensa en contra del vendedor. ¿Alteraría esta consideración la conclusión de que la prohibición de tales disposiciones impone a los consumidores costos financieros más elevados? No por necesidad, ya que la obligación de la indemnización elevaría los costos de los vendedores, y tratarían de dirigir esos costos a sus clientes.

[43] *Cf.* Williams *vs.* Walker-Thomas Furniture Co., 350 F.2d, 445 (D.C. Cir., 1965). El ejemplo es extremo; negar al comprador toda parte del monto de la reventa de la mercancía recogida se consideraría un castigo y, por lo tanto, estaría prohibido. Véase después, en IV.14, un análisis más amplio de la recuperación del bien.

una cantidad escasa de dinero salen beneficiados de los términos "duros" cuando la alternativa sería el pago de precios más elevados. Sin embargo, los beneficios pueden ser ilusorios, en su gran mayoría (véase después XIV.3), y es probable que liberar a los consumidores de los términos duros corrija el desbalance creado por las leyes del bienestar, que alientan los préstamos riesgosos truncando el aspecto negativo del riesgo.[44]

La coacción suele discutirse en términos de "poder de negociación", una ventaja transaccional creada por la posesión de un monopolio, ya sea del tipo que involucra los intereses de los agentes antimonopólicos, o del tipo bilateral, temporal o situacional. El término no debe utilizarse para significar la oferta de un contrato sobre la base de tómalo o déjalo; pero se puede utilizar, de manera inofensiva si bien no revestida de interés alguno, para señalar que una de las partes de la transacción tiene mucho mejores alternativas que la otra parte, aunque no sea a causa del monopolio. Por ejemplo, si hay muchos más hombres que mujeres en edad de contraer matrimonio, las mujeres serán en promedio más capaces de obtener mejores oportunidades matrimoniales que los hombres. En este sentido del poder de negociación, en una situación de escasez los vendedores tienen mayor poder de negociación que los compradores, y en una situación de abundancia excesiva los compradores tienen mayor poder de negociación que los vendedores.

Más de la mitad de los estados tienen leyes que prohíben la "especulación de precios", definida con frecuencia como el cobro de precios "inconscientemente" altos. Estas leyes rara vez se aplican. Sin embargo, la fuerte escalada de los precios de la gasolina en 2005, causada por el huracán Katrina y también por el huracán Rita, que siguió casi inmediatamente (los huracanes impidieron la importación de petróleo crudo y provocaron el cierre temporal de una serie de refinerías que se encontraban en su trayectoria), provocó una oleada de amenazas al cumplimiento, la imposición de algunas multas y propuestas para que la legislación federal prohibiera los aumentos "inconscientemente excesivos" en los precios de la gasolina. Es probable que la reducción inesperada de la oferta de un bien genere beneficios anormales para los proveedores del bien. La reducción disminuye la producción, lo que se traduce en un aumento de precio a medida que los consumidores ofertan unos contra otros por la producción reducida. Y es probable que la reducción de la producción disminuya los costos unitarios de los vendedores porque éstos suelen vender en un rango de producción en el que sus costos van en aumento (si disminuyeran, los vendedores tendrían un incentivo para aumentar más la producción). Cuando ocurren tanto el aumento del precio como la caída del costo, los beneficios tienden a dispararse al alza. Se sabe que los beneficios de algunas gasolineras aumentaron en 400% poco después de que el

[44] Véase Eric A. Posner, "Contract Laws in the Welfare State: A Defense of the Unconscionability Doctrine, Usury Law, and Related Limitations on the Freedom to Contract", 24 *J. Legal Stud.*, 283 (1995).

huracán Katrina tocara tierra. En tiempos de catástrofe, que involucran a consumidores perjudicados, el espectáculo de los vendedores que se benefician de la angustia de los consumidores, a la vez que (al parecer) aumentan su angustia al cobrarles precios altos, es una fuente de gran resentimiento.

Pero, ¿es éste un caso de coacción? Los refinadores y los distribuidores que aumentaron los precios después de que los huracanes interrumpieron la refinación de gasolina no crearon la situación que dio lugar a la reducción de la oferta. (Si lo hubieran hecho, digamos, al acordar el aumento de precios por encima del nivel actual, habrían sido castigados por violar las leyes antimonopolio.) En el caso Alaska Packers' no había una escasez preexistente de mano de obra que justificara la exigencia de los marineros de salarios más altos; los marineros *crearon* la escasez, al negarse a trabajar. Debido a que la escasez de gasolina provocada por el huracán que elevó los precios no era una escasez artificial, sino natural, sancionar a los proveedores de gasolina por cobrar un precio de equilibrio de mercado habría creado una escasez. Los consumidores aún habrían estado pagando un precio más alto que antes de la escasez, pero lo habrían hecho en forma de tiempo de espera en la fila de la gasolinera o de movilidad restringida (si manejaban menos debido a la escasez). Aquellos consumidores que necesitaban más la gasolina, al no poder expresar su necesidad presentando una oferta mejor que la de otros consumidores para el suministro limitado, habrían padecido más la escasez. Los beneficiarios habrían sido las personas con costos de tiempo menores y sin demanda urgente de gasolina.

IV.10. Principios fundamentales de los daños contractuales

Cuando se comprueba un incumplimiento contractual se plantea el problema de encontrar el remedio apropiado. Hay una diversidad de posibilidades que pueden ordenarse como sigue, aproximadamente de acuerdo con su severidad creciente:

1. La pérdida de confianza del promisario (los costos en los que incurrió por tener una confianza razonable en que el promitente cumpliría el contrato).
2. La pérdida de esperanza (la pérdida de la ganancia esperada del contrato); los daños líquidos (daños efectivamente especificados en el contrato como el remedio monetario por un incumplimiento).
3. Los daños consecuenciales (efectos de ondas del incumplimiento sobre la empresa del promisario).
4. La restitución (al promisario de las ganancias obtenidas por el promitente a causa de su incumplimiento).
5. El cumplimiento específico (la orden para que el promitente cumpla o se le impondrá un castigo por desacato al tribunal), y

6. Un castigo (a diferencia de los daños líquidos, que son una forma de daños compensatorios) especificado en el contrato, u otros daños punitivos.

Al decidir cuál remedio se otorgará, que el incumplimiento sea o no oportunista hace la diferencia. Si un promitente incumple su promesa sólo para aprovecharse de la vulnerabilidad de un promisario en un contexto (el contexto contractual normal) en el que el cumplimiento es secuencial antes que simultáneo, también podríamos aplicarlo en contra del promitente. Supongamos que A paga a B por adelantado los bienes y, en lugar de entregarlos, B emplea el dinero para construirse una alberca. En tal caso, la restitución es un remedio atractivo. Podemos impedir esta clase de comportamiento volviéndolo inútil para el promitente, lo que hacemos obligándolo a entregar al promisario todas sus ganancias derivadas del incumplimiento; ninguna sanción más ligera lo disuadiría. (¿Por qué no hacer que su conducta también se considere como delito o que sea delito? Porque, como veremos en el capítulo VII, las sanciones penales deben reservarse para situaciones en las que las sanciones monetarias son inadecuadas.)

La mayoría de los incumplimientos de contrato no son oportunistas. Muchos son involuntarios; el cumplimiento resulta imposible a un costo razonable… quizás a cualquier costo. Este punto tiene relación con la función aseguradora de los contratos. Otros incumplimientos son voluntarios, pero (como veremos enseguida) eficientes, lo que desde un punto de vista económico es el mismo caso que el de un incumplimiento involuntario. Estas observaciones explican la centralidad de los remedios para el derecho contractual (los remedios impuestos por hechos que pueden ser eficientes deben diseñarse cuidadosamente para evitar la disuasión excesiva) y le dan la razón a Holmes cuando dijo que no es la política del derecho obligar a que se respeten los contratos, sino sólo requerir que cada parte escoja entre el cumplimiento en los términos del contrato y la compensación a la otra parte por cualquier pérdida resultante del incumplimiento.[45]

Esta aseveración contiene una intuición económica importante. En muchos casos resulta antieconómico inducir el cumplimiento total de un contrato después de que se ha roto. Acepto comprar 100 000 adminículos elaborados a la orden como componentes de una máquina que yo fabrico. Después de haber recibido 10 000, se derrumba el mercado de mi máquina. Notifico inmediatamente a mi proveedor que estoy rescindiendo el contrato, y admito que mi rescisión es un incumplimiento. Cuando se notificó la rescisión, mi proveedor no había iniciado todavía la fabricación de los otros 90 000 adminículos, pero me informa que se propone completar su cumplimiento en los tér-

[45] Oliver Wendell Holmes, "The Path of the Law", 10 *Harv. L. Rev.,* 457, 462 (1897) ("la obligación de cumplir un contrato en el derecho común significa un pronóstico de que debes pagar daños si no puedes cumplirlo, y nada más").

minos del contrato y cobrarme consiguientemente. Los adminículos hechos a la orden no tienen ningún uso fuera de mi máquina, y poseen un valor de desecho insignificante. Si se le otorgara al proveedor un remedio que lo indujera a completar el contrato después del incumplimiento, se desperdiciarían recursos valiosos. La ley está alerta ante este peligro y, bajo la doctrina de la mitigación de los daños, no otorgaría al proveedor el pago de daños por ningún costo en el que incurriera al continuar la producción tras el aviso de la rescisión.

¿Pero no es irreal el peligro si el Teorema de Coase es cierto? Sólo hay dos partes, y hay un conjunto de precios para que el proveedor se olvide de insistir en sus derechos contractuales, lo que mejorará la situación de ambas partes. Pero éste es justamente otro ejemplo del monopolio bilateral; los costos de transacción serán elevados *porque* sólo hay dos partes.

Supongamos que se han producido y entregado todos los adminículos al comprador, quien se niega a pagar el precio del contrato. Se permitirá al vendedor demandar por ese precio. Una vez que el comprador tiene los bienes, la reventa es más barata para él que para el vendedor, por lo que el derecho lo hace aceptar y pagar por ellos. Si el vendedor tiene todavía los bienes, es probable que la reventa sea más barata para él, lo cual hará porque su único derecho es el pago de daños: la diferencia entre el precio del contrato y el precio al que puede vender los bienes a otro comprador.

Supongamos que el contrato de los adminículos es rescindido por el vendedor antes que por el comprador. Yo necesito realmente esos 100 000 adminículos sobre pedido para mi máquina, pero el proveedor, después de producir 50 000, se ve obligado a suspender la producción. Otros proveedores están en posibilidad de proveer los adminículos restantes que yo necesito, pero insisto en que el proveedor original complete su cumplimiento del contrato. Si la ley obliga a la terminación, el proveedor tendrá que hacer arreglos con otros productores a fin de completar su contrato conmigo. Quizá sea más costoso para él conseguir un proveedor alternativo que para mí conseguirlo directamente (después de todo, yo conozco mejor que nadie mis propias necesidades); de otro modo, el proveedor original lo habría hecho voluntariamente a fin de minimizar su responsabilidad por el incumplimiento. De nuevo, obligar a la terminación del cumplimiento del contrato (o a negociaciones costosas para liberar al promitente) generaría un desperdicio de recursos, y de nuevo la ley limita a la víctima al mero pago de daños simples.

¿Pero qué *son* los daños contractuales simples? El objetivo del otorgamiento de un incentivo al promitente para que cumpla su promesa, a menos que el resultado fuese un uso ineficiente de los recursos, puede lograrse entregando al promisario su ganancia esperada de la transacción. Si el proveedor del primer ejemplo (el del incumplimiento del comprador) recibe su ganancia esperada de la producción de 10 000 adminículos, no tendrá ningún incentivo para fabricar los 90 000 no deseados. No queremos que los fabrique; nadie los

quiere. En el segundo ejemplo (el del incumplimiento del vendedor), si yo recibo mi ganancia esperada del trato con el proveedor original, me vuelvo indiferente acerca de que complete o no su cumplimiento.

En estos ejemplos el incumplimiento fue cometido sólo para evitar una pérdida mayor, pero en algunos casos una de las partes se ve tentada a incumplir su contrato simplemente porque su ganancia derivada del incumplimiento superaría a su ganancia derivada de la terminación del cumplimiento del contrato. Si la ganancia superara también a la ganancia esperada por la otra parte de la terminación del cumplimiento del contrato y, por lo tanto, también al pago de daños por el incumplimiento, habría un incentivo para cometer un incumplimiento. Por lo tanto, en este caso el otorgamiento del pago de daños no disuadirá a alguien de incumplir el contrato; no debería. Es un incumplimiento eficiente. Supongamos que firmo un contrato para entregar 100 000 adminículos sobre pedido, a 10 centavos la pieza, para el uso de A en su fábrica de calentadores de agua. Tras haber entregado 10 000, se me presenta B y me explica que necesita de inmediato 25 000 adminículos sobre pedido porque de otro modo se verá obligado a cerrar su fábrica de pianolas a gran costo, y ofrece pagarme 15 centavos por cada pieza. Le vendo los adminículos y en consecuencia no cumplo a tiempo la entrega a A, causándole una pérdida de 1 000 dólares en ganancias. Habiendo obtenido una ganancia adicional de 1 250 dólares por la venta a B, me encuentro en una situación mejor aún después de rembolsarle su pérdida a A, y B también se encuentra en mejor situación. El incumplimiento es superior en términos de Pareto. Es cierto que, si me hubiese negado a venderle a B, éste podría haber acudido a A y negociado una cesión a su favor de una parte del contrato que A tiene conmigo. Pero esto habría introducido un paso adicional, con costos de transacción adicionales, y más elevados porque sería una negociación de monopolio bilateral.

¿No podría haberse evitado el peligro de una desistencia de los incumplimientos eficientes de contrato simplemente redefiniendo el concepto legal del incumplimiento de contratos de tal modo que sólo las rescisiones ineficientes contaran como incumplimientos? No. Recordemos que una función importante de los contratos es la asignación de los riesgos a los mejores tomadores de riesgos. Si el riesgo se materializa, la parte a quien se le asignó deberá pagar. No es más relevante el hecho de que esa parte no hubiera podido impedir la ocurrencia del riesgo a un costo razonable, quizás a cualquier costo, que el hecho de que una compañía de seguros no hubiera podido impedir el incendio que destruyó el edificio que aseguró. El incumplimiento de contrato corresponde a la ocurrencia del evento contra el cual se contrató el seguro.

Consideremos ahora el caso en el que la pérdida de la esperanza, es decir, la pérdida de la ganancia esperada del contrato, supera a la pérdida de la confianza: el gasto en que incurrió la víctima del incumplimiento al cumplir con su parte del contrato. Un fabricante acepta venderme una máquina por

100 000 dólares, para entregar a seis meses. El día siguiente a la firma del contrato el fabricante se retracta, advirtiendo que perdería 5 000 dólares al precio del contrato. La pérdida de la confianza del comprador —la suma de los costos en que ha incurrido irreparablemente a resultas del contrato— es cero, pero supongamos que le costaría 112 000 dólares obtener otra máquina, por lo que su pérdida esperada es de 12 000. ¿Por qué debiera permitírsele que insista en una medida de los daños que le otorga más (12 000 dólares adicionales) que lo que ha perdido efectivamente? Porque el otorgamiento de la pérdida de la confianza alentaría los incumplimientos ineficientes. La ganancia neta para el comprador del cumplimiento contractual, 12 000, sería mayor (en 7 000 dólares) que la pérdida neta del vendedor, así que al hacer de esa ganancia neta esperada un costo del incumplimiento para el vendedor se desalienta un cumplimiento ineficiente.

¿Qué ocurrirá si la pérdida de la confianza supera a la pérdida de la esperanza? En Groves *vs.* John Wunder Co.,[46] el demandado, como parte de un trato más amplio, había aceptado nivelar un terreno propiedad del demandante, y caprichosamente incumplió su compromiso. El costo de la nivelación habría sido de 60 000 dólares, y el valor de la tierra, después de la nivelación, no sería de más de 12 000 dólares; la Gran Depresión del decenio de 1930 se había iniciado después de la firma del contrato. El tribunal otorgó al demandante 60 000 dólares razonando que tenía derecho a obtener el cumplimiento que había contratado y que al demandado no le importaba si su cumplimiento hubiera hecho o no más valiosa la propiedad del demandante, o en cuánto. Sin embargo, no era un caso, conocido para nosotros por nuestra discusión de la compensación justa en el capítulo anterior, en el que el valor y el precio de mercado fuesen diferentes. La tierra en cuestión era una parcela comercial, y si el demandante hubiese deseado el cumplimiento en lugar de los 60 000 dólares, habría ejercitado una acción de cumplimiento específico. No ejercitó tal acción y, lo que es más revelador aún, no empleó el dinero que ganó del demandado para nivelar la tierra.[47] La medida de los daños que utilizó el tribunal, si la hubiese sabido el demandado desde el principio, lo habría vuelto indiferente entre el rompimiento de su promesa de nivelar la tierra y su cumplimiento, mientras que la eficiencia dictaba el incumplimiento: los 60 000 dólares que valdrían la mano de obra y los materiales que se hubieran consumido en la nivelación de la tierra habrían comprado menos que un aumento de 12 000 dólares en su valor.[48]

[46] 286 N.W. 235 (Minn., 1939).

[47] John P. Dawson, William Burnett Harvey y Stanley D. Henderson, *Cases and Comment on Contracts,* 18-19 (8ª ed., 2003).

[48] ¿Habría nivelado en efecto la tierra el demandado, en los términos del contrato, si hubiese sabido por adelantado cuál sería la medida de los daños en caso de que no lo hiciera? No, porque habría sabido que, a cualquier precio acordado entre 12 000 y 60 000, el demandante habría preferido que no se llegara a un acuerdo por la tierra. ¿Se deduce de aquí que el economista de-

Es cierto que si no se hubiera impuesto el cumplimiento del contrato se habría dado al demandado una ganancia inesperada. Pero el cumplimiento forzoso del contrato dio al demandante una ganancia inesperada igual y opuesta: un colchón, que casi seguramente no habían intentado las partes, contra el impacto de la Depresión sobre los valores de la tierra. Dado que el demandante, como propietario de la tierra, y no el demandado, un contratista, habría disfrutado el beneficio de cualquier aumento inesperado del valor de la tierra, si las partes hubiesen meditado sobre el asunto habrían deseado probablemente que el demandante soportara también la carga de cualquier baja inesperada de ese valor.

La medida de la esperanza podría ser una aproximación mejor, que la medida de la confianza, a la pérdida económica real de la víctima, además de producir mejores incentivos. En el equilibrio competitivo a largo plazo, los ingresos totales de los vendedores de un mercado son exactamente iguales a sus costos totales; no hay "ganancia" en un sentido económico, sino sólo el rembolso de los costos del capital, del esfuerzo empresarial y de otros insumos, con inclusión de los esfuerzos de comercialización que condujeron al contrato. Todos estos elementos del costo se excluyen por la medida de la confianza de los daños, lo que tenderá a subestimar, por lo tanto, los costos sociales del incumplimiento. Aunque el incumplimiento ocurra antes de que la víctima haya empezado a cumplir, es posible que la víctima haya incurrido en costos (especialmente los costos de la búsqueda antes del contrato, pero éstos no son recuperables como daños: es difícil que la otra parte los calcule con antelación y su asignación a un contrato específico suele ser difícil). Supongamos que la víctima no ha incurrido en gastos antes de empezar a cumplir, de modo que el costo de la confianza es cero hasta que se inicie el cumplimiento. Entonces, si los costos de la confianza fuesen la única medida de los daños, las partes podrían rescindir sus contratos siempre que fuesen todavía ejecutorios. Excepto en situaciones especiales, no está claro cuál sería la ganancia social derivada de tal periodo de "enfriamiento" y podría haber una pérdida social a resultas de la incertidumbre y del aumento de las transacciones.

La medida de la esperanza no es perfecta. Al otorgar a la parte que cumple una ganancia garantizada, por decirlo así, en lo que en el caso habitual será una empresa más o menos riesgosa —es decir, garantizada si la otra parte incumple el contrato—, la medida de la esperanza puede inducir una confianza excesiva en la parte que cumple, justo como cualquiera otra forma del

biera ser indiferente a la medida de los daños en este caso? No, de nueva cuenta, porque la regla del pago de daños en Groves creó una situación de monopolio bilateral, en la que el demandado tendría un incentivo para amenazar con nivelar la tierra con el fin de reducir el precio del arreglo y el demandante tendría un incentivo para insistir en que se nivelara la tierra para elevar el precio del arreglo tan cercano a 60 000 como fuera posible. Afortunadamente, se dictó un fallo opuesto al de Groves en Peevyhouse *vs.* Garland Coal & Mining Co., 382 P.2d 109 (Okla., 1963).

seguro comercial tenderá a inducir al asegurado a relajar sus esfuerzos para evitar el azar contra el que se asegura. En principio, este problema podría resolverse mediante el "antiseguro", por el que las partes convendrían en que, en caso de un incumplimiento, el promitente pagaría los daños a un tercero (la compañía de "antiseguro") antes que al promisario. Esto daría al promisario el incentivo correcto para economizar en la confianza y ayudar al cumplimiento del promitente en lugar de obstruirlo. Como una compensación por el otorgamiento de su derecho al pago de daños a la compañía de antiseguro, el promisario recibiría de esa compañía, por adelantado, un pago igual al valor esperado de ese derecho.[49] Sin embargo, calcular ese valor esperado a menudo sería difícil, y traer a un tercero a escena complicaría aún más las cosas. Exigir que el pago de los daños de confianza sea razonable y permitir las cláusulas de daños líquidos en los contratos, un requisito y un permiso que forman parte de la legislación vigente, pueden ser alternativas adecuadas al antiseguro.

La aplicación de la medida de la esperanza implica también otros problemas. Comparemos los casos siguientes: *1)* Un inquilino no paga, y el arrendador renta la propiedad de inmediato a otro inquilino por una renta sólo ligeramente menor que la del inquilino que no pagó. En un juicio contra el inquilino que no pagó, por la renta adeudada sobre el resto del contrato de arrendamiento, ¿debiera exigirse al arrendador que deduzca la renta del inquilino sustituto? *2)* Un fabricante recibe de X un pedido de 1 000 adminículos, pero X se niega a aceptar la entrega y el fabricante revende los adminículos a Y a un precio sólo ligeramente menor que el que X había aceptado pagar. En un juicio contra X por las ganancias que se perdieron en la venta, ¿debiera exigirse al fabricante que dedujera las ganancias que recibió en la venta sustituta a Y?

La ley responde afirmativamente en el primer caso y negativamente en el segundo, y estas respuestas son correctas desde un punto de vista económico. La producción del arrendador es fija a corto plazo; no puede añadir un cuarto porque una familia más quiera rentárselo. La renta que recibe del inquilino sustituto en el primer caso es así una ganancia *hecha posible* por el incumplimiento del contrato del primer inquilino,[50] de modo que su pérdida verdadera es la diferencia existente entre las dos rentas.[51] Pero de ordinario un fabricante puede variar ligeramente su producción, aun a corto plazo. El incumplimiento de X no hizo posible que el fabricante obtuviera una ganan-

[49] Véase Robert Cooter y Ariel Porat, "Anti-Insurance", 31 *J. Legal Stud.*, 203 (2002).

[50] Adviértase que para que esta regla del pago de daños sea eficiente es esencial un requerimiento de la mitigación de los daños, ya que de lo contrario el inquilino no tendrá ningún incentivo para encontrar un inquilino sustituto a menos que el inquilino que no pagó se declare insolvente.

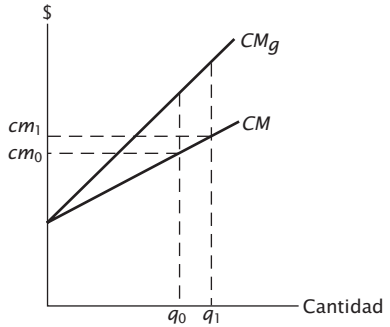
[51] ¿Depende esto de que todos los demás departamentos estén ocupados o no? Sí, porque si existen vacantes el arrendador no remplaza al inquilino que se va con otro; podía haber rentado al nuevo inquilino uno de los departamentos desocupados. Por lo tanto, la salida del antiguo inquilino no hizo posible otra renta.

cia vendiendo a Y; porque si X no hubiera incumplido, el fabricante podría haber proveído todavía a Y de 1 000 adminículos. La ganancia de la venta a Y es una ganancia que el fabricante habría obtenido independientemente del incumplimiento, de modo que su pérdida verdadera es la ganancia total esperada de la venta a X.

Sin embargo, la ganancia puede ser difícil de estimar. Es la diferencia existente entre el precio del contrato y el costo del cumplimiento del vendedor, ¿pero cuál es el costo? Deben distinguirse dos componentes: los costos fijos y los costos variables. Los costos fijos (llamados a veces "costos generales", lo que es equívoco, como veremos enseguida) no varían con la cantidad producida; los costos variables sí lo hacen. Virtualmente todos los costos son variables a largo plazo, pero en el caso de un contrato que deberá cumplirse en un tiempo relativamente breve puede omitirse el largo plazo. Y, a corto plazo, costos como la renta, el seguro, los impuestos (algunos), los sueldos de los funcionarios y los intereses de la deuda a largo plazo son fijos. Dado que no se ahorrarán si el vendedor no hace esta venta particular, estos costos no debieran restarse del precio del contrato para determinar cuánto empeora la situación del vendedor a resultas del incumplimiento. En concreto, supongamos que el precio del contrato es de 11 dólares por adminículo, los costos variables del vendedor son seis dólares, y sus costos fijos son cuatro dólares. El incumplimiento costará al vendedor cinco dólares (el precio del contrato menos los costos variables), no un dólar, porque todavía incurrirá en los costos fijos de cuatro dólares que asignó a la venta para fines contables. Pero debemos tener cuidado de que los costos fijos sean *realmente* fijos. Supongamos que si el contrato no se hubiera celebrado el vendedor habría cerrado, y de este modo habría eliminado todos sus costos (esto supone que no tendría ninguna obligación residual por concepto de impuestos, pensiones, rentas o intereses). Entonces, todo lo que el vendedor hubiera perdido por el incumplimiento habría sido un dólar; por la misma razón, todo lo que hubiera ganado con la terminación del contrato habría sido un dólar. En este ejemplo, los costos generales como los sueldos son realmente costos variables.[52]

A fin de medir los costos variables del cumplimiento de un contrato, nos vemos tentados a dividir simplemente los costos totales de la compañía (tras deducir sus costos fijos) por su producción total para obtener un costo variable medio, y suponer que ése es el costo en que habría incurrido el vendedor al fabricar los adminículos adicionales que se requieren para cumplir con el contrato. Pero es probable que le cueste más al vendedor dicha fabricación. Pregúntese por qué el vendedor no fabrica más adminículos de los que produce. La respuesta más probable es que la producción de una cantidad mayor lo llevaría a una situación en la que encontraría las deseconomías de escala, lo que elevaría sus costos unitarios por encima del precio de mercado.

[52] Véase Autotrol Corp. *vs.* Continental Water Systems Corp., 918 F.2d, 689 (7th Cir., 1990).



GRÁFICA IV.1

Por ejemplo, es posible que tuviera que contratar trabajadores adicionales, y en tal caso podría tener que pagar salarios más altos para alejarlos de otros productores. Y podría incurrir en estos costos mayores no sólo en su producción adicional sino en toda su producción: si ofreciera salarios mayores a los trabajadores nuevos a fin de arrebatarlos de las manos de sus empleadores actuales, podría tener que ofrecer los mismos salarios mayores a sus trabajadores actuales si el pago de un salario mayor a los nuevos empleados aumentara el nivel salarial general en el mercado. Debido a tales consecuencias del aumento en la producción, un contrato aparentemente lucrativo podría darle al vendedor sólo una ganancia moderada.

Esto se representa en la gráfica IV.1. Se supone que el vendedor tiene costos variables solamente, de modo que *CM* es su lista de costos variables medios y de costos totales medios. *CM*g es la lista de los costos marginales del vendedor. Si el costo marginal es creciente, como en la gráfica IV.1, el costo variable medio será creciente también, pero con mayor lentitud, porque el aumento de los costos totales al pasar a producciones sucesivamente mayores está siendo promediado en toda la producción de la empresa. La venta perdida habría aumentado la producción de la empresa de q_0 a q_1, lo que haría aumentar sus costos medios de cm_0 a cm_1. Si los daños causados por la venta perdida se determinan restando del precio del contrato $(q_1 - q_0) \times cm_0$ (una cifra conocida), antes que $(q_1 - q_0) \times cm_1$ (una cifra hipotética, basada en una venta que no se realizó), la empresa será compensada en exceso por el incumplimiento. (El costo marginal y otros conceptos se discuten más extensamente en los capítulos IX y XIII.) Pero el cambio del costo generado por un contrato nuevo tenderá a ser pequeño cuando el contrato contempla un aumento proporcionalmente pequeño de la producción de la empresa (en la gráfica IV.1 la proporción es grande). Por lo que toca a tales contratos, el costo variable medio de la empresa antes del cumplimiento del contrato es probablemente una buena aproximación al costo efectivo del cumplimiento del contrato.

Todo lo que se dice en esta sección acerca de las "ganancias" perdidas podría sugerir a algunos lectores que todos los vendedores son monopolistas. Una empresa competitiva fija su precio igual a su costo marginal; no tiene ninguna "ganancia" que perder si fracasa una de sus ventas. Lo cierto es que, fuera de algunos mercados muy organizados, como los de productos agrícolas y de instrumentos financieros, en general una empresa podrá elevar su precio ligeramente por encima de su costo marginal sin que sus ventas desaparezcan por completo, debido a una ventaja de ubicación o a un producto que es especialmente atractivo para algunos compradores. El estado de cosas en el que un mercado contiene muchas empresas, con una pequeña cantidad de poder monopólico cada una, en el sentido de que puede maximizar sus ganancias mediante el cobro de un precio superior a su costo marginal, se llama "competencia monopólica".

Pero incluso las empresas en un mercado plenamente competitivo, antes que monopolísticamente competitivo, tienen costos fijos y, por lo tanto, deben cobrar un precio superior a su costo marginal. La medida de expectativa de los daños no presupone una competencia monopolística, sino sólo que la empresa tiene algunos costos fijos. La medida de expectativa mide la contribución que habría hecho la venta a los costos fijos si se hubiera realizado. Pero, si el vendedor es un monopolista, su precio de venta *contendrá* un rendimiento monopólico, y en tal caso puede argüirse que la medida de la esperanza es demasiado generosa porque puede inducir al comprador a pasar por alto un incumplimiento eficiente, ya que al decidir sobre el cumplimiento del contrato el comprador no comparará el costo social del incumplimiento con sus costos de cumplimiento del contrato, sino con un potencial fallo de pago de daños que contenga rentas monopólicas.[53]

Veamos ahora un problema con los daños de los compradores. A contrata con B la venta de 1 000 adminículos para entrega el 11 de enero de 2006, a un precio de un dólar por adminículo. El 11 de julio de 2005, A le dice a B que no podrá venderle los adminículos. La retractación anticipada es un incumplimiento de contrato y faculta a B para buscar los adminículos en otra parte. (Vimos antes la retractación anticipada en el caso de los adminículos hipotéticos.) El día del incumplimiento, el precio de los adminículos es de dos dólares. B podría acudir al mercado y "cubrirse" firmando otro contrato a futuro (es decir, un contrato para entrega futura, el 11 de enero de 2008) por 1 000 adminículos, y entonces sus daños serían de 1 000 dólares, suponiendo que los precios corrientes y los futuros fueran iguales, dos dólares, en la fecha del cubrimiento. En lugar de ello, B espera hasta el 11 de enero y ese

[53] David D. Friedman, "An Economic Analysis of Alternative Damage Rules for Breach of Contract", 32 *J. Law & Econ.,* 281 (1989). (Vimos antes un argumento similar en III.6, referente a la incautación de embarcaciones por parte del gobierno en tiempos de guerra.) Como veremos en el capítulo IX, la confrontación de los compradores con alternativas (socialmente) falsas constituye una objeción económica central contra el monopolio.

día compra 1 000 adminículos para entrega inmediata al precio entonces vigente de tres dólares por adminículo, perdiendo en consecuencia 2 000 dólares. ¿Debiera estar facultado para recuperar la pérdida mayor? No debiera estarlo (la respuesta en términos de la economía y también en términos de la ley).[54] Pues supongamos que el precio del mercado hubiese disminuido entre julio y enero por debajo del precio del contrato, digamos a 50 centavos. Al permitir que B espere, se le hace posible apostar sin ningún riesgo para él. Por mucho que aumente el precio de los adminículos, B está protegido porque el contrato fija un tope a su gasto. Pero si el precio baja más allá del precio del contrato, B se queda con la diferencia. No era ésta la intención de las partes. Su contrato desplazaba todo el riesgo de los cambios de precios, no sólo el riesgo de un aumento del precio, de B a A al fijar el precio de los adminículos para entrega futura.

Un contrato a futuro es un buen ejemplo de la función de los contratos de desplazar el riesgo, pero un ejemplo mejor aún es el contrato de futuros.[55] Un contrato de futuros tiene la misma forma general que un contrato a futuro, excepto que normalmente no se contempla la entrega (sólo de 1 a 2% de todos los contratos de futuros se traduce en una entrega efectiva). Los contratos de futuros se emplean generalmente en conexión con los productos agrícolas y los metales, cuyos precios son volátiles.[56] Supongamos que un silo tiene un acervo de granos que sus clientes no necesitarán durante los próximos seis meses. Si el silo no desea soportar el riesgo de las fluctuaciones de precios durante este periodo, puede vender un contrato de futuros que prometa la entrega a un precio fijo dentro de seis meses (digamos a tres dólares el bushel). Cuando transcurran los seis meses, el silo venderá el grano al precio del mercado, que será ahora, digamos, de sólo un dólar el bushel, y al mismo tiempo cancelará el contrato de futuros comprando un contrato idéntico para entrega inmediata a un dólar. La última transacción le dará una ganancia neta de dos dólares por bushel (la ganancia en su contrato de futuros), que contrarrestará exactamente su pérdida por tener que vender su grano a sólo un dólar por bushel antes que a tres dólares.

La ventaja del contrato de futuros sobre el contrato a futuro es que un vendedor no necesita, para protegerse contra una baja del precio, buscar y encontrar a alguien que desee efectivamente comprometerse a recibir la entrega de sus bienes dentro de seis meses. Sólo tiene que encontrar a alguien

[54] Thomas H. Jackson, "'Anticipatory Repudiation' and the Temporal Element in Contract Law: An Economic Inquiry into Contract Damages in Cases of Prospective Nonperformance", 31 *Stan. L. Rev*., 69 (1978); pero véase Alan Schwartz y Robert E. Scott, *Commercial Transactions: Principles and Policies,* 323-325 (1982).

[55] Véanse Dennis W. Carlton, "Futures Market: Their Purpose, Their History, Their Growth, Their Successes and Failures", 4 *J. Futures Mkts.,* 237 (1984); Lester G. Telser, "Why There Are Organized Futures Markets", 24 *J. Law & Econ.,* 1 (1983).

[56] Los contratos financieros de futuros tienen una importancia creciente, pero plantean problemas diferentes, de la clase que se discute en el capítulo XVI.

que piense que el precio aumentará. Así pues, los contratos de futuros amplían el campo de la especulación. Y la especulación facilita la protección y, al otorgar a ciertos individuos (los especuladores) un interés en el pronóstico correcto de los precios, aunque no estén implicados en la producción o el consumo del bien comerciado en el mercado, aumenta la cantidad de la información existente en el mercado acerca de los precios. *Cf.* antes III.5.

Hoy en día muchos contratos de futuros son financieros antes que futuros de productos básicos: son especulaciones sobre los precios futuros de divisas, acciones, bonos y otros instrumentos financieros. Son parte del mundo de los valores derivados y los discutiremos en el capítulo XVI, donde veremos que las permutas de cobertura por incumplimiento crediticio resultan ser un contrato especulativo, así como una forma de seguro de crédito.

IV.11. Daños consecuenciales

Una cuestión importante de los daños de los compradores es el hecho de que el comprador puede obtener o no lo que se llama pago de daños *consecuenciales* por el incumplimiento del vendedor. (La probabilidad de que los compradores incurran en el daño consecuencial es mayor que la de los vendedores, ya que, por lo general, éstos pueden vender su producto a alguien más si fracasa la venta prevista inicialmente.) El epítome de la hostilidad tradicional del derecho común a los daños consecuenciales en los casos contractuales es el caso famoso de Hadley *vs.* Baxendale.[57] Consideremos la variante siguiente de los hechos de ese caso. Un fotógrafo comercial compra un rollo de película para tomar fotografías de los Himalayas para una revista. El costo del revelado de la película se incluye en el precio de compra de la película. El fotógrafo incurre en fuertes gastos (incluida la contratación de un avión) a fin de realizar la tarea. Envía la película por correo al fabricante, pero se extravía en el cuarto de revelado y jamás aparece.

Comparemos los efectos de incentivo del hecho de permitir que el fotógrafo recupere daños por las consecuencias plenas del incumplimiento con los de limitar su recuperación al precio de la película. La primera alternativa crea pocos incentivos para evitar pérdidas similares en el futuro. El fotógrafo no tomará ninguna precaución: le dará igual optar por la terminación exitosa de su asignación que por recibir la compensación plena por su fracaso. (Ésta es una pequeña exageración porque desconoce sus costos de litigación.) Es probable que tampoco el fabricante de la película tome precauciones adicionales, porque no puede identificar las películas cuya pérdida sería en extremo costosa, y, a menos que haya muchas de ellas, quizá no convenga tomar precauciones adicionales con *todas* las películas que revele. En cambio, la se-

[57] 1.9 Ex. 341, 156 Eng. Rep. 145 (1854).

gunda alternativa induciría al fotógrafo a tomar precauciones que resulten a la vez asequibles y eficaces: el empleo de dos rollos de película o la petición de un manejo especial cuando envíe la película al revelado.

El principio general es que, si un riesgo de pérdida es conocido sólo por una de las partes del contrato, la otra parte no es responsable de la pérdida que ocurra. Este principio induce a la parte que conoce el riesgo a tomar precauciones apropiadas, o, si cree que la otra parte podría ser quien impide o dispersa (el asegurador) la pérdida con mayor eficiencia, a revelar el riesgo a la otra parte y pagarle por asumir el riesgo. Así se crean incentivos para asignar el riesgo en la forma más eficiente.[58] Sin embargo, existen motivos para pensar que incluso si el vendedor es quien soporta mejor el riesgo en cuestión, el comprador puede sentir temor de revelarle el riesgo al vendedor y pedirle que lo asuma porque la información transmitida acerca de la vulnerabilidad de aquél podría ocasionar que éste cobrara un precio más alto.[59]

La regla de Hadley *vs.* Baxendale a menudo se enuncia en términos de lo "previsible" que sea el daño consecuencial del comprador. Pero esta palabra, aunque sea tan favorecida en el derecho, es enloquecedoramente vaga. Volvamos al caso del fotógrafo comercial: aunque el revelador de la película desconoce las consecuencias de extraviarla o echarla a perder, sí sabe que tales pérdidas pueden ocurrir. Es sólo que el comprador, aunque no puede evitar el incumplimiento mismo, se encuentra en mejor posición que él para evitar las consecuencias del incumplimiento. La regla de Hadley *vs.* Baxendale se asemeja así a la regla del derecho aplicable a los cuasidelitos (no aceptada universalmente) de que no se permitirá a la víctima de un accidente recuperar los daños que podría haber evitado ajustándose el cinturón. La víctima no podría haber evitado el accidente obrando de ese modo, así que la responsabilidad del victimario no es afectada; pero la extensión de su lesión sí lo es.[60]

No se aplica la regla de Hadley *vs.* Baxendale en los casos en los que el daño consecuencial consiste en las ganancias perdidas a resultas del incumplimiento. Supongamos que le ofrezco 140 000 dólares por su casa que tiene un valor de mercado de 150 000 dólares; usted acepta la oferta, pero después rescinde el contrato, y yo lo demando por 10 000 dólares, mi ganancia perdida.

[58] Adviértase la relación analógica entre el caso de Hadley *vs.* Baxendale y la regla que gobierna la liberación de un contrato en razón de la imposibilidad (antes, IV.5): Hadley crea una regla de descarga parcial para los casos en los que el promisario es quien soporta mejor un porcentaje del riesgo que genera una falta de cumplimiento.

[59] Véase Louis E. Wolcher, "Price Discrimination and Inefficient Risk Allocation under the Rule of *Hadley vs. Baxendale*", 12 *Research in Law & Econ.*, 9 (1989). Éste es un ejemplo de los problemas estratégicos que ocurren en la negociación de los contratos, un tema de interés particular para los teóricos de juegos.

[60] *Cf.* EVRA Corp. *vs.* Swiss Bank Corp., 673 F.2d, 951 (7th Cir., 1982), en que se aplica este principio ("las consecuencias evitables") a la cuestión de si un banco debiera ser responsable de los daños consecuenciales derivados de una falla negligente en la transferencia de fondos electrónicos.

No se le permitiría defenderse sobre la base de que no había ninguna razón para que pensara que la transacción era tan rentable para mí. De otro modo sería difícil que un buen negociador recibiera el pago de daños, a menos que antes de la firma del contrato hubiera hecho revelaciones que redujeran la ventaja de ser un buen negociador: revelaciones que le impedirían al comprador apropiarse las ganancias derivadas de sus esfuerzos por identificar un recurso que estaba subvaluado en su uso actual. Ésta es justamente la aplicación del principio de Laidlaw *vs.* Organ en un contexto de daños.

IV.12. Castigos, daños punitivos, daños líquidos, confiscaciones y reglas de no modificación

En algunas ocasiones, un contrato especificará los daños que se otorgarán si hay un incumplimiento (tal como una demora), y si la especificación es una estimación razonable *ex ante* de los daños probablemente derivados del incumplimiento, se hará cumplir bajo el rubro de los "daños líquidos", aunque los daños efectivos resulten ser menores o mayores. Pero, si es evidente desde el principio que la especificación está diseñada para otorgar a la víctima del incumplimiento mucho más de lo que podría esperar perder efectivamente a resultas del incumplimiento, entonces será una *cláusula de castigo* y no se hará cumplir. El nombre está bien escogido; la imposición a un violador de un costo en exceso del daño efectivamente causado por la violación es la esencia de una sanción penal, como veremos en el capítulo VII.

Podría parecer obvio que la ley no haría cumplir —y en efecto la ley estadunidense no lo hace— las cláusulas de castigo en los contratos. Un castigo impediría los incumplimientos eficientes, al igual que los ineficientes, al hacer el costo del incumplimiento para quien rompe el contrato mayor que el costo del incumplimiento para la víctima; también crearía problemas de monopolio bilateral. Un castigo crea una gama de negociaciones; el promitente pierde más del incumplimiento (el castigo) que lo que gana la víctima del cumplimiento (el valor neto del contrato para ella, equivalente a los daños compensatorios simples), de modo que cualquier punto entre estas dos cifras es el precio por excusar el incumplimiento que hará que ambas partes se encuentren en una situación mejor. Asimismo, el castigo podría inducir a la víctima potencial a provocar un incumplimiento, ya que se beneficiaría de ello. Éstas son buenas razones para no otorgar pagos de daños punitivos por los incumplimientos de contrato, al menos en general (más adelante se discuten las excepciones posibles). Pero no son buenas razones para negarse a hacer cumplir forzosamente las cláusulas de castigo negociadas voluntariamente; éstas no se pondrían en los contratos —no con frecuencia, por lo menos—, a menos que las partes esperaran que las ganancias superaran a los costos que acabamos de identificar. Y podría ocurrir así. Supongamos que yo sé que cumpliré

mis contratos, pero me resulta difícil convencer a otros de este hecho. Al firmar una cláusula de castigo, comunico una información creíble acerca de mi propia estimación de mi confiabilidad, una información útil para determinar en cuáles términos hacer negocios conmigo.[61]

Otra razón para una cláusula de castigo es la de compensar al vendedor por un gran riesgo de impago. Supongamos que los compradores que no pagan serán a menudo insolventes o por otra razón incapaces de cubrir al vendedor sus daños plenos. Entonces la recuperación "inesperada" de un castigo en algunos casos permitirá, al contrarrestar las pérdidas sufridas en otros, que los vendedores tomen riesgos mayores y cobren precios menores. Además, una cláusula de castigo tiene sentido económico si hay alguna probabilidad de que un incumplimiento pase inadvertido, por ejemplo, si el vendedor pudiera culpar plausiblemente al comprador del fracaso del producto del vendedor para desempeñarse como se intentaba. Si no hay una cláusula de castigo, la responsabilidad del vendedor por los daños esperados será menor que los daños efectivamente sufridos por el comprador, y habrá demasiados incumplimientos.

Una explicación económica, aunque muy especulativa, de la renuencia de los tribunales a ejecutar las cláusulas de castigo es que éstas podrían conducir a un aumento del número de juicios contractuales sin ninguna ganancia claramente compensatoria. Esta sugerencia podría parecer paradójica, ya que entre mayor sea la sanción por un incumplimiento de contrato menos probable será un incumplimiento. Pero en la medida en que las cláusulas de castigo permitan que se celebren contratos con personas que de otro modo se hallan en un riesgo de impago demasiado grande para ser partes contratantes atractivas, la desistencia incrementada del incumplimiento podría contrarrestarse con un mayor número de contratos.[62] El problema del "azar o riesgo moral" (el promisario tiene más que ganar del incumplimiento que del cumplimiento cuando hay una cláusula de castigo en el contrato, de modo que podría tratar de provocar un incumplimiento) podría contrarrestar también el efecto de una cláusula de castigo para la desistencia de los incumplimientos.

Además, una cláusula de castigo desalienta las modificaciones y renegociaciones que impiden el incumplimiento.[63] Dado que la responsabilidad contractual es una responsabilidad estricta, un promitente puede cometer un incum-

[61] Aquí presentamos un ejemplo aún más dramático del uso del señalamiento para establecer la credibilidad. "Las creencias en los Valhallas que recompensan a los soldados valerosos muertos en batalla, o los infiernos que castigan a los cobardes, a los ladrones y a los mentirosos, fueron legitimadas y hechas más visibles por el gran desperdicio regio en los funerales, testimonio de las creencias de los líderes en una vida en el más allá." Donald T. Campbell, "Legal and Primary Group Social Controls", 5 *J. Social Biological Structures,* 431, 437 (1982).

[62] Un argumento similar, que encontraremos en el capítulo XV, es que el otorgamiento de mejores remedios para los acreedores en las quiebras podría aumentar en efecto el número de las quiebras al envalentonar a los acreedores para que presten dinero a prestatarios menos dignos de crédito.

[63] Tai-Yeong Chung, "On the Social Optimality of Liquidated Damage Clauses; An Economic

plimiento de contrato simplemente porque está incapacitado para cumplir su promesa, no porque no quiera hacerlo, y sin embargo no tiene la defensa de la imposibilidad ni la de *force majeure*. Muy a menudo, en tal caso las partes aceptarán modificar el contrato, con una compensación apropiada para el promisario. La modificación es menos probable cuando hay una cláusula de castigo, porque esa cláusula aumenta el precio que demandará el promisario por permitir que el promitente se libere de la férula del contrato original. Si fuese muy costoso un juicio para hacer efectiva tal cláusula (suponiendo que tales cláusulas fuesen legalmente exigibles) el promisario se vería presionado para buscar un arreglo. Pero si el cumplimiento forzoso no es costoso (y no debería serlo, ya que la cláusula especifica una suma cierta como el monto de los daños por un incumplimiento), el promisario podría insistir en busca de algo muy cercano al castigo, y el promitente podría concluir que quizá le convendría dejar que el promisario demandara. Como veremos en el capítulo XXII, el litigio puede ser muy costoso para el sistema judicial, aunque no lo sea para las partes —esto es, las partes podrían externalizar algunos de los costos de su litigio—. A fin de limitar la externalidad, los tribunales podrían negarse a hacer cumplir las disposiciones contractuales que hagan más probable el litigio, a menos que haya fuertes eficiencias compensatorias.

Además, los castigos aumentan el riesgo de la quiebra derivada del impago contractual y, al hacerlo, incrementan el número y, por ende, el costo total de las quiebras (los costos de recursos, no simplemente las transferencias pecuniarias), y una parte de ese costo es externo a las partes, como veremos cuando discutamos la quiebra en el capítulo XV. Un argumento relacionado, pero macroeconómico, en contra de las cláusulas de castigo es que, si fuesen comunes, se incrementaría la amplitud del ciclo económico haciendo aún mayor de lo que es el número de las quiebras en las depresiones y recesiones.[64]

Se ha argumentado que no hay cláusulas de castigo *reales*, que vista *ex ante*, y dejando de lado los casos de fraude o coacción u otras fallas patológicas en el proceso de contratación, una cláusula de castigo sólo compensa al promisario por un riesgo de pérdida poco común, tal como una probabilidad por arriba del promedio de impago de un promitente de alto riesgo o el riesgo de un impago indetectable.[65] Esto no es cierto en el caso del señalamiento; el prestatario que desea convencer al prestamista de que es digno de crédito podría aceptar pagar un castigo que superara los costos esperados del prestamista, sólo para establecer la credibilidad del prestatario. Sin embargo, en la mayoría de los casos es cierto que la cláusula de castigo no compensa excesi-

Analysis", 8 *J. Law, Econ & Organization,* 280 (1992); Eric L. Talley, "Contract Renegotiation, Mechanism Design, and the Liquidated Damages Rule", 46 *Stan. L. Rev*., 1195 (1994).

[64] Daniel A. Farber, "Contract Law and Modern Economic Theory", 78 *Nw. U. L. Rev.,* 303, 335 (1983).

[65] Alan Schwartz, "The Myth That Promisees Prefer Supracompensatory Remedies: An Analysis of Contracting for Damage Measures", 100 *Yale L. J*., 369 (1990).

vamente *ex ante*. Los problemas que hemos discutido con las cláusulas de castigo surgen porque compensan en exceso *ex post*. Si tales problemas aumentan las tasas de litigios podrían explicar, y concebiblemente incluso justificar, la hostilidad de los tribunales hacia las cláusulas de castigo.

Sin embargo, existen tres argumentos más a favor de la aplicabilidad de las cláusulas de castigo que pueden inclinar la balanza en contra de la regla del derecho común que las prohíbe: *1)* los contratos deben cumplirse según sus términos a menos que existan objeciones económicas convincentes, y no solamente conjeturales; *2)* los sistemas de derecho civil imponen cláusulas de castigo (aunque se reservan el derecho a reducir el castigo, pero no hasta llegar a la pérdida real, es decir, el nivel de los daños compensatorios, en caso de que sea extremadamente excesivo) sin que se caiga el cielo,[66] y *3)* obligar al cumplimiento de las cláusulas de castigo eliminaría el tener que diferenciar los castigos de los daños líquidos, lo cual puede ser difícil. Consideremos las cláusulas de "tómalo o paga", comunes en la industria del gas: requieren que el comprador tome la cantidad de gas que ha convenido en comprar o pague el precio de esa cantidad de gas aunque tome menos, en lugar de pagar simplemente los daños del vendedor, que serían la diferencia entre ese precio y el precio al que el vendedor podría revender el gas a otro comprador. ¿No es ésta una compensación excesiva? No necesariamente. Un contrato para el suministro de gas implica a menudo fuertes costos fijos para el proveedor, tales como la construcción de un gasoducto hasta las instalaciones del comprador. Los costos fijos no se pueden evitar por una disminución de la cantidad vendida a resultas del incumplimiento del comprador. Si los costos fijos representan un alto porcentaje de los costos totales, la cláusula de tómalo o paga podría no ser una sobrecompensación significativa para el vendedor.

Podría parecer que la ley debiera requerir que todo contrato contenga una cláusula de liquidación de daños según la teoría de que las partes saben mejor que los tribunales cuáles serán probablemente los daños. Pero eso sería un error. Los costos de la estimación de los daños pueden ser menores cuando los daños surgen mucho tiempo antes, cuando el contrato se firma. Y bajo el régimen de las cláusulas obligatorias de liquidación de daños se incurriría en tales costos en todas las negociaciones contractuales, no sólo en la pequeña fracción de los casos en los que se rompe un contrato y sigue un juicio.

Una regla relacionada con la prohibición de las cláusulas de castigo es la prohibición de la concesión de daños punitivos por incumplimiento de contrato. La regla tiene sentido, pues recordemos que el incumplimiento de contrato rara vez se puede ocultar; la responsabilidad por el incumplimiento es estricta, por lo que muchos incumplimientos son involuntarios y no intencionales, e incluso los incumplimientos deliberados pueden ser eficientes y, por lo tanto, socialmente deseables antes que ilícitos; por lo general, los da-

[66] Aristides N. Hatzis, "Having the Cake and Eating It Too: Efficient Penalty Clauses in Common and Civil Contract Law", 22 *Int'l Rev. Law & Econ.,* 381 (2003).

ños compensatorios se pueden calcular con un grado de precisión razonable; y las víctimas probables de los incumplimientos de contratos pequeños pueden compensarse en el precio del contrato por no tener un buen remedio legal para el incumplimiento. Sin embargo, se ha argumentado que se deben conceder los daños punitivos en los casos de incumplimiento *intencional*, porque revela que el promitente es una persona "mala" a la que se debe castigar.[67] Pero la otra parte no será "mala" mientras se compense en su totalidad a la víctima del incumplimiento; simplemente está ejerciendo la opción del incumplimiento que le concede el derecho contractual, según Holmes. (Más adelante retomaremos la teoría de la opción del contrato de Holmes en IV.16.)

La excepción principal a la negativa de la ley para la concesión de daños punitivos incluso para los incumplimientos "intencionales" del contrato es la negativa de mala fe de una compañía de seguros para cumplir el contrato de seguro. Por "mala fe" nos referimos a que es francamente infundada; y si es infundada el asegurado tiene una demanda sólida para los beneficios del seguro y puede exigirlos con altas posibilidades de ganar. ¿Por qué entonces debiera permitírsele obtener daños punitivos por añadidura? Decir que "los incumplimientos eficientes no existen en el contexto de los seguros, [y que] los incumplimientos de contratos de seguros sólo pueden ser oportunistas"[68] no es una buena respuesta; suele haber una verdadera incertidumbre fáctica o jurídica acerca del deber de la aseguradora de indemnizar al asegurado. En algunas jurisdicciones, el resultado de la regla de mala fe es que las aseguradoras que no se consideran responsables de la política presentan demandas de juicio declarativo para demostrar su no responsabilidad. Una justificación mejor de la regla, aunque sólo en algunos casos, es que el asegurado puede ser altamente dependiente de la pronta recepción de los beneficios del seguro (por ejemplo, si su casa se ha incendiado), y esto podría permitir que la compañía de seguros impusiera condiciones duras en las negociaciones de arreglos. Pero este problema podría manejarse, al igual que otros incumplimientos de contrato oportunistas, mediante el principio de la coacción, sin necesidad de otorgar daños punitivos.

Otra táctica evasiva en el derecho contractual que los tribunales permiten es la de demandar a nuestro médico por el cuasidelito de mala práctica profesional antes que sólo demandar por incumplimiento del contrato con él de parte del demandante. La razón puede ser que el derecho aplicable a los cuasidelitos se orienta hacia los casos que involucran una lesión física, y el derecho contractual no. Sin embargo, esta consideración no es aplicable cuando se hace una demanda por mala práctica en contra de nuestro abogado más que contra nuestro médico. En ese caso, la razón dada para el remedio de los cuaside-

[67] Oren Bar-Gil y Omri Ben-Shahar, "An Information Theory of Willful Breach", 107 *Mich. L. Rev.,* 1479 (2009).

[68] Dominick C. Capozzola, "Note, First-Party Had Faith: The Search for a Uniform Standard of Culpability", 52 *Hastings L. J.*, 181, 196 (2000).

litos es que el abogado es un fiduciario de su cliente más que un vendedor de servicios, y el incumplimiento de una obligación fiduciaria es un cuasidelito. Sería más directo decir que una obligación fiduciaria es un término implícito de algunos contratos, más que utilizar la presencia de tal obligación para modificar una disputa surgida de un contrato derivándola de un caso de contrato a un caso de cuasidelito.

El derecho común, al mismo tiempo que prohíbe las cláusulas de castigo y los daños punitivos, permite que los vendedores guarden depósitos y pagos de abonos, aunque el resultado sea el de otorgar al vendedor más dinero que cualquier estimación razonable de sus daños. El Código Comercial Uniforme ha cambiado esta regla de tolerancia de las confiscaciones referente a las ventas de bienes; empero, la regla antigua persiste en otros casos, sobre todo en los casos que impliquen a la tierra, si bien con una limitación importante: un acreedor hipotecario que lleva a cabo procedimientos de secuestro no se convierte por ello en dueño de la propiedad. Si la propiedad se vende en un remate, el deudor tiene derecho a recibir el excedente (si hubiera alguno) del precio de venta sobre la suma adeudada por la hipoteca.

¿Por qué se trata con más benignidad a las confiscaciones que a los castigos? Para ejecutar una confiscación no se requiere una demanda; por lo tanto, la aplicación de este remedio es más barata para el sistema legal; la parte sujeta a la confiscación suele ser el pagador; es menos probable que la parte que paga cometa un incumplimiento involuntario que la parte que cumple (y no queremos que se nos impongan castigos por incumplimiento involuntario); y es poco probable que una confiscación sea ruinosa, ya que se limita al dinero que ya pagó la parte en contra de quien se solicita la confiscación.

El derecho común se rehúsa a hacer obligatoria la cláusula del contrato que prohíbe que las partes lo modifiquen. (Puede parecer que esto no tiene relación con los castigos y las confiscaciones, pero continúe leyendo.) La "teoría" es que las partes siempre pueden acordar la eliminación de la cláusula de no modificación y, habiéndolo hecho, acordar la modificación del contrato. Pero sólo pueden acordar hacer esto si la ley prohíbe una cláusula de no modificación vinculante. Existen buenas razones por las que las partes podrían desear prohibir la modificación: para protegerse contra la coacción (recordemos la forma en que los marinos utilizaron la coacción en el caso Alaska Packers' para obligar a la modificación de su contrato de trabajo, aunque el tribunal deshizo la modificación). Sin embargo, podrían surgir problemas graves en los casos en los que la falta de modificación impusiera costos devastadores, si se prohibiera la recontratación para permitir la modificación.

IV.13. El cumplimiento específico

Cuando un tribunal decreta el cumplimiento específico, esto significa que ordena a la parte que ha roto el contrato que lo cumpla, so pena de ser declarada en desacato del tribunal si no lo hace.[69] Por lógico que pueda parecer este remedio, es excepcional en el derecho anglonorteamericano. El promisario debe, como sugiere la sentencia de Holmes citada, conformarse con los daños[70] —a menos que los daños sean difíciles de computar porque no hay en el mercado buenos sustitutos para el cumplimiento del contrato por parte de quien lo rompe—. La excepción se ha tragado generalmente a la regla en el caso de los contratos de venta de inmuebles. Supongamos que tengo un contrato para comprar una casa y que el vendedor incumple. La estimación de los daños puede ser muy difícil porque, como hemos visto (en III.4), es posible que yo valúe la casa más que el mercado. Si se remite al comprador al pago de daños en tales casos, podría generarse una subestimación sistemática de los costos de un incumplimiento, ya que un tribunal se guiará necesariamente por el precio de mercado y se mostrará escéptico ante la afirmación interesada e incomprobable de un comprador sobre el valor idiosincrásico de la casa para él.[71] Además, el precio de mercado puede ser difícil de determinar porque los edificios no son un producto remplazable y mueble; el precio de venta de una casa a menudo no es una indicación definitiva del valor de mercado de otra casa, similar en apariencia.

El problema de valuación se evita decretando el cumplimiento específico, pero se crea otro problema. El hecho de que el vendedor incumpliera puede indicar que hay otra transacción que tiene valor en mayor medida del que tendría la terminación de una venta contractualmente convenida; en tal caso, los resultados del decreto de un cumplimiento específico no son catastróficos, ya que el vendedor siempre puede pagarle al comprador para que renuncie al derecho del cumplimiento específico, y presumiblemente lo hará si

[69] Hay una bibliografía abundante sobre la economía del cumplimiento específico. Véanse, por ejemplo, Anthony T. Kronman, "Specific Performance", 45 *U. Chi. L. Rev.,* 351 (1978); Alan Schwartz, "The Case for Specific Performance", 89 *Yale L. J.,* 271 (1979); William Bishop, "The Choice of Remedy for Breach of Contract", 14 *J. Legal Stud.,* 299 (1985); Paul G. Mahoney, "Contract Remedies and Options Pricing", 24 *J. Legal Stud.,* 139 (1994).

[70] Ésta es una regla general de los remedios equitativos. El remedio equitativo más común es una orden, y una orden de cumplimiento específico de un contrato es simplemente un tipo de orden obligatoria (una orden para que alguien haga algo en lugar de abstenerse de hacer algo). La discusión en el texto puede generalizarse a la mayoría de los remedios equitativos.

[71] Adviértase la tensión entre la regla que permite el cumplimiento específico como algo natural en los casos que involucran la tierra y la regla que otorga al gobierno las facultades del dominio eminente. La primera regla únicamente muestra esa sensibilidad a la diferencia potencial entre el mercado y los valores verdaderos que no existe en las reglas de dominio eminente que se analizaron en el capítulo III.

una transferencia sustituta genera un precio mayor. Pero la negociación adicional tendrá algún costo. De nuevo nos encontramos ante un monopolio bilateral, con una gama de negociación delimitada por los costos del incumplimiento para con el promisario en el límite inferior y por los costos de cumplimiento por parte del promitente en el límite superior.[72] Pero esto no es un problema cuando se utiliza el cumplimiento específico para obligar a un propietario a transferir la propiedad de acuerdo con un contrato de venta. No existe el riesgo de que el cumplimiento específico, si se aplica, impondrá costos de producción elevados para cumplir con el contrato, porque el contrato no requiere producir nada. Ésta puede ser otra razón por la que se ordena el cumplimiento específico de manera rutinaria en los casos de bienes inmuebles.[73]

Otra razón para preferir los remedios de daños al cumplimiento específico en general es que, para el tribunal, el juicio de daños es un asunto de una sola vez; el tribunal emite un fallo, y el alguacil va y vende algún activo del demandado si éste se niega a pagar voluntariamente la condena. Pero el cumplimiento específico, como otros remedios equitativos, requiere que el tribunal mantenga el caso en sus archivos hasta que se complete el cumplimiento, a fin de que, en caso necesario, pueda responder al argumento del demandante de que el demandado no está cumpliendo de buena fe.

IV.14. Autoayuda, condiciones contractuales y cumplimiento sustancial

Supongamos que un consumidor compra un automóvil en abonos, toma posesión, pero no completa sus pagos. La suma adeudada podría ser demasiado pequeña para justificar los costos del juicio, aunque los tribunales menores —utilizados más por los vendedores que por los compradores, adviértase— facilitan el pago de las reclamaciones contractuales pequeñas. En lugar de eso, el vendedor podría tomar el automóvil, venderlo y remitir el dinero de la venta al comprador, menos los costos de la venta y la suma que se deba al vendedor en los términos del contrato. Si se permitiera que el vendedor se quedara con todo el dinero, esto sería una confiscación que ya no está permitida.

¿Debiera importar el hecho de que el vendedor revenda el automóvil al menudeo o al mayoreo? Podría parecer que, dado que los precios al menudeo son mayores que los precios al mayoreo, el vendedor debiera estar obligado a vender al menudeo para favorecer al comprador. Pero la ley no re-

[72] Sin embargo, si se trata de un caso de imposibilidad física en la que el promitente, después de haber asumido el riesgo de la imposibilidad, es responsable del incumplimiento, no se puede ordenar el cumplimiento específico y, en consecuencia, se tendría que idear algún sustituto monetario. En tal caso, el promitente es, en realidad, un asegurador.

[73] Véase Steven Shavell, "Specific Performance *versus* Damages for Breach of Contract: An Economic Analysis", 84 *Tex. L. Rev.*, 831 (2006).

quiere esto, y con razón. Dado que la venta al menudeo y al mayoreo son industrias competitivas, tanto el precio de venta al menudeo como al mayoreo deben ser iguales al costo de la venta del bien. El precio al menudeo será mayor sólo porque los costos al menudeo son mayores que los costos al mayoreo. El pago neto al dueño del bien debe ser el mismo.

El derecho de recuperación de la posesión puede enunciarse en el lenguaje de las "condiciones", que es una parte importante del derecho contractual. El derecho del comprador para obligar al vendedor a cumplir con el contrato entregándole el automóvil está condicionado a que le pague al vendedor el precio total del automóvil. Si el comprador incumple esta parte del contrato, el vendedor tiene derecho a rescindir su cumplimiento y recuperar la posesión del automóvil. Supongamos que el comprador hubiera incumplido antes de la entrega del automóvil. Entonces el vendedor podría haber invocado la doctrina de las condiciones como una razón para negar la entrega. Pero no podría haberse quedado con el dinero que el comprador le había pagado ya al celebrar el contrato de venta. Los pagos podrían exceder grandemente a los daños sufridos por el vendedor a causa del incumplimiento, en cuyo caso el remedio de la autoayuda podría impedir los incumplimientos eficientes.

Sin embargo, lo común es que la autoayuda del *comprador*[74] no cause problemas. B pide adminículos a A, pero cuando se reciben e inspeccionan advierte que son defectuosos y los regresa. Éste es un remedio más barato, en términos sociales y privados, que el hecho de que B se quede con los adminículos y demande a A por daños, o incluso los venda por cuenta de A. Éste presumiblemente tiene un mejor sentido de lo que debe hacerse con los adminículos defectuosos que B (repararlos, desecharlos, venderlos a alguien como de "segunda clase", etcétera).

Aun así, ¿es consistente este enfoque con el remedio del vendedor cuando el comprador es quien incumple (véase antes IV.10)? Allí vimos que se permite que el vendedor demande por el precio del contrato una vez que ha entregado los bienes al comprador, obligando en realidad al comprador a comprarlos, bajo la teoría de que una vez que éste tiene los bienes es probable que para él sea más barato revenderlos (suponiendo que no se los quiera quedar) que para el vendedor. La diferencia entre eso y el caso de incumplimiento por parte del vendedor es que en este último no queremos imponer al comprador la carga de revender los bienes (para mitigar sus daños) y demandar al vendedor por sus daños, suponiendo que esto incluye los costos de transacción de la reventa en que incurrió. En cambio, permitimos que el comprador devuelva los bienes; sus daños suelen ser nulos o mínimos, y es probable que los costos totales de transacción involucrados en el remedio del incumplimiento no sean menores.

Si B no realiza la inspección de los bienes cuando llegan y en consecuencia no descubre los defectos sino mucho tiempo después, podría considerar-

[74] Véase George L. Priest, "Breach and Remedy for the Tender of Nonconforming Goods under the Uniform Commercial Code: An Economic Approach", 91 *Harv. L. Rev.,* 960 (1978).

se que ha aceptado los bienes y tendrá que pagar por ellos. Éste podría parecer un requerimiento inútil, ya que la ley permite que B demande a A por incumplimiento de la garantía en la provisión de bienes defectuosos a pesar de la aceptación. Pero entre más se demore B en devolver los bienes, más se habrán depreciado éstos; es decir, más costoso será el remedio de la autoayuda del comprador para el vendedor.

Supongamos que contrata la construcción de una casa, y cuando llega el día en que el constructor debe entregarle la casa terminada advierte que ésta no cumple todas las especificaciones del contrato. ¿Debiera estar facultado para rechazar la entrega, aunque no haya pagado nada del precio de la casa? De nuevo, la cuestión fundamental es la relación existente entre el costo de su remedio de autoayuda para el constructor y el pago de daños a que tendrá usted derecho si en cambio ha demandado. Supongamos que, en virtud de que la casa fue diseñada a la medida de sus gustos peculiares, su valor de reventa es muy bajo; el constructor incurrirá en una pérdida de 50 000 dólares si usted tiene derecho a rescindir el contrato, mientras que los daños a los que tendría usted derecho sólo ascenderían a 1 000 dólares. En este caso, la autoayuda no es un remedio eficiente. Por lo tanto, pronosticamos (y descubrimos) que la ley no permite al promisario que excuse su falta de pago en razón de un incumplimiento menor por parte del promitente.[75] La ley compara implícitamente el costo de la autoayuda del promisario para el promitente con los daños del promisario y se niega a permitir la autoayuda cuando el primero de estos costos supera sustancialmente al segundo. Así pues, podemos esperar que se permita la autoayuda con mayor frecuencia en los casos de bienes muebles y remplazables que en los casos de bienes que se fabrican sobre pedido.

Sin embargo, si *hubiera* un incumplimiento en el caso de la construcción, ¿cuál será el remedio? En Ruxley Electronics & Construction Ltd. *vs.* Forsyth,[76] el demandado había convenido en contratar la construcción de una alberca para los demandantes que tendría 2.3 metros de profundidad en su punto más hondo. Después de construida la alberca, los demandantes descubrieron que sólo tenía 1.9 metros en su punto más hondo y sólo 1.8 metros en el punto en el que la gente se tiraría clavados en la alberca. Se otorgó el pago de ciertos daños a los demandantes, pero menor que el costo de la reconstrucción de la alberca a la profundidad contractualmente especificada, que era lo que habían buscado. Ese costo habría sido desproporcionado a la disminución del valor de mercado de su propiedad a resultas del incumplimiento. En efecto, se descubrió que tal disminución era cero, un poco implausiblemente, pero quizás el hecho de que la alberca menos profunda fuese más barata de operar y mantener habría contrarrestado para la mayoría de los compradores la pérdida de placer en su uso.

[75] Véase un ejemplo famoso en Jacob & Youngs, Inc. *vs.* Kent, 129 N.E. 889 (N.Y., 1921) (Cardozo, J.).

[76] [1995] 3 All E.R. 268 (H.L.).

La pérdida de placer no se había discutido en Jacob & Youngs.[77] El incumplimiento contractual había sido la instalación de una tubería de agua de marca equivocada. La tubería estaba oculta en las paredes y la marca equivocada era tan buena como la que se había especificado en el contrato; por lo tanto, los daños eran realmente nulos. Pero ¿podría el comprador haber argumentado a favor de otro remedio? El contrato puede haber especificado una marca en particular con el fin de limitar el criterio del constructor, debido a la preocupación de que cualquier sustitución deliberada podría ser de una marca de menor calidad.[78] O el comprador podría haber querido la marca especificada, y no otra, por alguna razón peculiar, que un tribunal no podría apreciar. Se puede decir que la restitución del beneficio del constructor derivada de la sustitución se habría justificado en ambas situaciones.[79]

Un contraargumento es que las partes entienden que los contratos de construcción otorgan libertad al constructor para llevar a cabo sustituciones en respuesta a los cambios en el precio y la disponibilidad de los distintos componentes, y que se compensa al comprador por la libertad del constructor mediante el cobro de un precio más bajo a cambio de dar flexibilidad al constructor.[80] En un contrato, las especificaciones son legalmente exigibles, por lo que el comprador tiene un remedio de daños contra una sustitución que, aprobada o no por el arquitecto (y puede que no haya un arquitecto), redujera el valor de la propiedad.

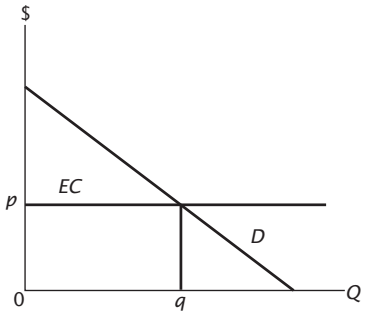
Aunque el otorgamiento de daños por la pérdida de beneficios o servicios en el caso Ruxley parece el punto medio feliz entre cero y el costo de la reconstrucción, la determinación de tales daños "subjetivos", como ya se advirtió, está llena de incertidumbre. Uno de los jueces en Ruxley empleó el concepto económico de "excedente del consumidor" (véase antes en I.2) para precisar el problema. Recordemos que el excedente del consumidor generado por la venta de un producto es el área debajo de la curva de demanda al precio de mercado menos el costo total (el precio multiplicado por la producción) para los consumidores. Si el producto se vende en condiciones competitivas, existe la probabilidad de que el excedente del consumidor sea muy grande, *EC* en la gráfica IV.2. Ésta es una medida del beneficio que obtienen

[77] Véase la nota 76.

[78] Véase Steve Thel y Peter Siegelman, "Willfulness *vs.* Expectation: A Promisor-Based Defense of Willful Breach Doctrine", 107 *Mich. L. Rev.*, 1517, 1527-1528 (2009).

[79] El mandato judicial, ordenar al constructor que remplace la tubería de las paredes con la tubería que se especifica en el contrato, sería una alternativa, pero el costo podría superar a la pérdida para el comprador derivada de no tener la tubería deseada.

[80] Como explica Todd D. Rakoff, "The Implied Terms of Contracts: Of 'Default Rules' and 'Situation-Sense'", en *Good Faith and Fault in Contract Law* 191, 209 (Jack Beatson y Daniel Friedmann, comps. 1995), la función de gran parte de los detalles de un contrato de construcción "no es la de expresar un propósito, o aun un 'capricho', del propietario, sino más bien la de servir como una comunicación del arquitecto al constructor; los detalles se modifican comúnmente bajo la autoridad del arquitecto durante todo el proceso de construcción".



GRÁFICA IV.2

los consumidores, como un grupo, del hecho de ser capaces de comprar el producto al precio competitivo. Si las albercas se venden en un mercado competitivo, los consumidores valuarán las albercas en más que su precio. Es decir, *la mayoría* de los consumidores lo harán. En la intersección de la demanda y el precio se encontrará el consumidor marginal, quien valúa la alberca en sólo la cantidad que le cuesta; éste no obtiene ningún excedente del consumidor.

Así pues, la alberca que deseaban los demandantes de Ruxley les habría conferido probablemente, aunque no con seguridad por la razón que acabamos de mencionar, un beneficio no pecuniario en exceso de su precio. Estos demandantes obtuvieron *una* alberca, de modo que la pérdida real era la diferencia entre el excedente del consumidor que la alberca deseada les habría conferido y el excedente del consumidor que les confería la alberca sustituta. Desafortunadamente, el excedente del consumidor es muy difícil de medir. Uno de sus parámetros esenciales, la curva de demanda, es una lista de precios hipotéticos antes que reales. Además, el excedente del consumidor es una agregación del valor subjetivo (descontado el precio) que obtienen del producto *todos* los consumidores. El agregado podría ser más fácil de determinar que el excedente obtenido por uno de los consumidores, el demandante en un juicio. La pendiente negativa de la curva de demanda sugiere que el monto del excedente podría variar grandemente entre los consumidores.

En un caso como el de Ruxley, hay un argumento para ofrecer al demandante una elección entre la estimación de su pérdida subjetiva hecha por el tribunal y el ahorro hecho por el contratista gracias al incumplimiento. Imaginemos un caso en el que la pérdida subjetiva "real", que no se puede conocer, es de 500 dólares, el ahorro logrado por el contratista por haber roto el contrato es de 400 dólares, y los daños que el tribunal otorgará por la pérdida subjetiva del demandante son sólo de 300 dólares. Un contratista que prevea correctamente el fallo de un juicio sabrá que puede "ganar" 100 dólares por el incumplimiento. Puede desalentarse la tentación de este contratista de incumplir su promesa si el tribunal ofrece al demandante la alternativa que he sugerido (400 dólares en el ejemplo). Ésa parece una solución mejor que el

remedio draconiano de otorgar al comprador el costo de la reconstrucción, un remedio que no toma en cuenta la inevitabilidad de incumplimientos menores de los contratos de construcción. Sin embargo, la ganancia del contratista es el segundo mejor remedio, ya que supone que no hay pérdida de excedente del consumidor. Es un remedio restitutorio diseñado para impedir el fraude, antes que para compensar plenamente a un comprador por una pérdida.

IV.15. Contratos implícitos

Un médico se topa con un extraño que yace inconsciente en la calle, lo atiende y luego demanda el pago de honorarios. ¿Tendrá derecho? La ley dice que sí. La antigua terminología hablaba de un contrato implícito entre el médico y el extraño por la asistencia médica. Esta idea ha sido atacada como una ficción, y los autores modernos prefieren basar el derecho del médico en el principio del enriquecimiento injusto. Este término huele a moralidad, pero los juicios se explican mejor en términos económicos. (El derecho utiliza lenguaje moral principalmente porque proporciona un vocabulario familiar en el cual se discuten las obligaciones y los derechos. Tomarlo en sentido literal en un análisis jurídico constituye una causa común de errores.) El concepto de un contrato implícito es una abreviatura útil de un enfoque económico; destaca este enfoque la continuidad entre las cuestiones de los contratos expresos y las cuestiones que se tratan ahora bajo el rubro del enriquecimiento injusto.[81]

En el caso del médico, los costos de una transacción voluntaria serían prohibitivos. La causa es la incapacidad. En otros casos podría tratarse del tiempo (por ejemplo, el extraño está consciente pero sangra profusamente y no hay tiempo para discutir los términos). En tales casos, la ley considera si, de no haber sido prohibitivos los costos, las partes se habrían puesto de acuerdo y, de ser así, cuáles habrían sido entonces (aproximadamente) los términos. Si un tribunal confía razonablemente en que habría habido una transacción y en que se puede saber cuáles habrían sido sus términos esenciales (que el médico realiza sus mejores esfuerzos y que el paciente paga los honorarios normales del médico por el tratamiento recibido), no vacilará en redactar un contrato entre las partes después del hecho. Adviértase la analogía con el salvamento marítimo.

¿Habría alguna diferencia si el demandado puede probar que es un científico cristiano, de modo que, si hubiese estado consciente, no habría contratado los servicios del médico? No debiera haberla, a menos que en los otros

[81] Véase una discusión más amplia en William M. Landes y Richard A. Posner, "Salvors, Finders, Good Samaritans, and Other Rescuers: An Economic Study of Law and Altruism", 7 *J. Legal Stud.*, 83 (1978); Saul Levmore, "Explaining Restitution", 71 *Va. L. Rev.*, 65 (1985); Robert A. Long, "Note, A Theory of Hypothetical Contract", 94 *Yale L. J.*, 415 (1984).

casos de tratamiento de personas inconscientes se otorgaran premios a los médicos para compensarlos por el riesgo de que la persona inconsciente no desee realmente sus servicios (y, por lo tanto, no se le haría pagar por ellos).

Pero supongamos ahora que un hombre se para bajo mi ventana, tocando bellamente el violín, y cuando ha terminado llama a mi puerta y pide un pago por sus esfuerzos. Aunque disfruté su ejecución, me niego a pagar cualquier cosa por ella. El tribunal negaría la reclamación de un pago hecha por el violinista —por razonable que pudiera parecer el pago— alegando que, aunque el violinista me confirió un beneficio (y no con la intención de que fuese gratuito), lo hizo oficiosamente. Traducido de la terminología jurídica a la económica, esto significa que el violinista confirió un beneficio no negociado en circunstancias en las que los costos de una negociación voluntaria habrían sido bajos. En tales casos la ley insiste en que se siga la ruta voluntaria, y se encuentra en terreno económico firme al hacerlo así. No hay aquí ningún problema de monopolio bilateral porque no hay una negociación.

¿Sería diferente si el violinista hubiese sido contratado por mi vecino y por error tocara bajo mi ventana? Quizá, ya que no se elude un mercado deliberado y se ha otorgado un beneficio. ¿Y si el violinista, en lugar de tocar su violín, pagó por error un abono de mi hipoteca?[82] Puesto que no hay problema alguno en la valoración del beneficio, se debe otorgar la restitución.

IV.16. La teoría de la opción del derecho contractual *vs*. la teoría moral

Muchos estudiantes del derecho contractual creen que, porque un contrato constituye una promesa, el romperlo, al menos si se hace de manera deliberada (recordemos que muchos incumplimientos de contrato son involuntarios), es un acto ilícito. Los que apoyan este punto de vista tienden a pensar que los remedios para el incumplimiento de un contrato son inadecuados.

Pero, ¿es éste un punto de vista sólido? Hemos visto que Oliver Wendell Holmes argumentó, en el espíritu de la economía positiva, que cuando firmamos un contrato en el cual prometemos un cumplimiento específico (el suministro de un producto o la prestación de un servicio) compramos una opción para cumplir o para pagar daños. La característica de la opción es pronunciada, sobre todo, cuando el contrato contiene una cláusula de daños líquidos. Estamos prometiendo que cumpliremos o que pagaremos la suma que se especifica en la cláusula. Pero el carácter opcional del cumplimiento de contrato es general: mientras paguemos los daños otorgados por el tribunal en la demanda del promisario por incumplimiento de contrato, sea que se especifiquen en la cláusula o que se calculen de acuerdo con los principios

[82] Véase después, en IV.16, una discusión más amplia del enriquecimiento injusto. Véase también Christopher T. Wonnell, "Unjust Enrichment and Quasi-Contracts", en *Contract Law and Economics*, 454 (Gerrit De Geest [comp.], 2011).

de daños contractuales, no se nos puede culpar por incumplimiento, incluso si fue deliberado —incluso si, por ejemplo, no cumplimos simplemente porque alguien nos ofreció más dinero por el producto o servicio que nos habíamos comprometido a proveer en el contrato, y que no teníamos la capacidad de proveer a ambos, al promisario y al cliente nuevo, el más necesitado—. *En realidad* no hemos roto la promesa, porque lo que prometimos (aunque el contrato no se haya redactado de esta forma) era o lo uno o lo otro: no el cumplimiento sino *ya sea* el cumplimiento *o* la compensación por el costo del incumplimiento para la otra parte del contrato.

De este modo, la teoría de la opción es una teoría de "no culpabilidad" del derecho contractual. Como tal, explica por qué se otorgan rara vez remedios tales como los daños punitivos y la indemnización en los casos de contratos. Generalmente, la indemnización se concede cuando la ley quiere disuadir al demandado de observar determinada conducta más que sólo hacer que internalice sus costos. Por ejemplo, en un caso de derechos de autor, es probable que el infractor sea el explotador más eficiente de la obra con derechos de autor, pero, si es así, al limitar al titular de los derechos de autor o *copyrights* a los daños compensatorios se crearía un régimen de licencias obligatorias. Cualquier individuo que se considerara el productor más eficiente podría infringir impunemente, tratando los daños que adeudara al propietario de los derechos de autor como un derecho de licencia. Al hacer que la infracción carezca de valor para el infractor, la indemnización de las ganancias de éste obliga al infractor potencial a negociar con el propietario del derecho de autor, manteniendo así el régimen de los derechos de propiedad de la ley de los derechos de autor. Sin embargo, otorgar la indemnización en los casos de contratos (salvo en los casos de rescisión, donde el objetivo es regresar a las partes a su posición precontractual, la cual podría requerir que una de las partes renuncie a un beneficio que recibió de la otra) crearía un derecho de propiedad en el cumplimiento de un contrato; lo que generaría un problema de monopolio bilateral. El promitente que tenía una oportunidad mayor no podría acogerse a ella (de manera más precisa, no podría obtener beneficio alguno si aprovechara la oportunidad) sin negociar con el promisario. El problema es menos grave en los casos de derechos de autor, donde suele haber infractores potenciales múltiples que, si se les impide conservar los beneficios derivados de la infracción, en su lugar competirán para obtener una licencia del propietario del *copyright*.

El hecho de que la víctima de un incumplimiento de contrato pueda obtener en ocasiones el cumplimiento específico, el remedio de mandato, no contradice la teoría de no culpabilidad del derecho contractual . El remedio de mandatos sólo es posible cuando el "remedio conforme a la ley", es decir, el pago de daños, no está disponible. Porque en tal caso la garantía de contrato pierde su carácter de o lo uno o lo otro; más que una promesa de cumplimiento o el pago de daños, es una promesa de cumplimiento o nada (porque

el pago de daños no está disponible), y ésa no es una opción que el promisario habría aceptado. En cambio, un derecho *general* al cumplimiento específico habría hecho efectivamente que el derecho contractual se concibiera con base en las fallas. Un tribunal no estaría dispuesto a ordenar a la parte incumplidora que cumpliera, bajo pena de desacato civil y penal y multas potencialmente astronómicas si no lo hiciera, sin antes tomar en cuenta si los costos del cumplimiento para ésta superarían los beneficios de la otra parte. Y un derecho general al cumplimiento específico frustraría algunos incumplimientos eficientes. Si A rompe su contrato con B para vender a C porque C pagará más que el daño para B derivado del incumplimiento (equivalente al pago de daños), el incumplimiento aumenta el producto social: la situación de B no es peor, y A y C se encuentran en mejores situaciones. Pero si se le otorga el derecho al cumplimiento específico a B, A no puede vender a C sin tener que pagarle a B para acordar la terminación del contrato entre ellos, creando una situación de monopolio bilateral.

La teoría de la opción del contrato o de las obligaciones contractuales implica que la responsabilidad por el incumplimiento de un contrato es estricta, es decir, que la víctima del incumplimiento no necesita probar la culpa de quien incumple el contrato (otra razón por la cual el cumplimiento específico no puede ser el remedio estándar para el incumplimiento), porque la teoría de la opción es una teoría de no culpabilidad. La promesa es el cumplimiento o el pago de daños, por lo que, si se opta por el incumplimiento, incluso si uno se ve impedido de cumplir por circunstancias que están fuera de su alcance, debe pagar los daños. No tendría sentido que se nos excusara sólo porque el costo del cumplimiento superaría los beneficios; eso haría fútil a la opción.

Dicho de otra manera, una opción tiene un componente de seguro. Si usted me promete ya sea el cumplimiento o la indemnización en lugar del cumplimiento, me está asegurando contra las consecuencias de su incumplimiento. Como explicó Holmes, "las consecuencias de una promesa vinculante en el derecho común no se ven afectadas por el grado de poder que posee el promitente sobre el evento prometido. [...] En el caso de una promesa vinculante de que lloverá mañana, el efecto jurídico inmediato de lo que hace el promitente es que asume el riesgo del evento, dentro de ciertos límites definidos, como algo entre él y el promisario. No hace más cuando se compromete a entregar un fardo de algodón".[83] Usted hará tal promesa, otorgarme esa opción, si es el asegurador más barato contra el riesgo de incumplimiento. La responsabilidad estricta del incumplimiento reduce los costos de transacción al optimizar la asunción de riesgos, la función cumplida en el ámbito de los cuasidelitos por el seguro de responsabilidad formal.

Un principio acordado del derecho común de los contratos es la obligación de la buena fe en el cumplimiento de un contrato, y puede parecer que

[83] O. W. Holmes, Jr., *The Common Law*, 299-300 (1881).

esa obligación introduzca un elemento moral en el derecho contractual. En realidad, es una respuesta al hecho señalado al principio de este capítulo de que el cumplimiento de las partes en un contrato no suele ser simultáneo y que, como resultado, una de las partes se encontrará con frecuencia a merced de la otra en el transcurso del cumplimiento: la otra parte tendrá un monopolio temporal. El hecho de que la ley utilice términos morales como la "buena fe" no significa que las obligaciones legales sean deberes morales. El derecho utiliza lenguaje moral, principalmente, como se señaló antes, porque es un vocabulario familiar en el cual discutir las obligaciones y los derechos.

Un ejemplo paralelo es la atención excesiva en la obligación de los mejores esfuerzos en los contratos de negociación exclusiva: el exclusivo otorga al distribuidor un monopolio del producto durante la vida del contrato, y la obligación de los mejores esfuerzos le impide explotarlo. La obligación de la víctima del incumplimiento de contrato de mitigar sus daños es explicable en términos monopólicos. El incumplimiento crea un monopolio bilateral: las partes preferirían solucionar la demanda de la víctima, pero sólo pueden hacerlo negociando entre ellas. A la víctima del incumplimiento le gustaría que se acumulara el pago de los daños en que aún no ha incurrido porque eso aumentaría la oferta máxima de arreglo de la otra parte. La obligación de mitigar los daños impide que la víctima explote su monopolio temporal otorgado por el contrato.

En resumen, la obligación contractual de "buena fe" legalmente exigible es sólo una obligación para evitar la explotación de la posición de monopolio temporal que la parte contratante obtendrá en ocasiones en el transcurso del cumplimiento. No hay una obligación *general* de buena fe en el derecho contractual y, por tanto, no es necesario salir de la economía para explicar el principio. Si ofrece un precio bajo por algún bien a su propietario, usted no está obligado a decirle que considera que el bien está subvalorado, que él no se da cuenta de su valor de mercado y usted sí. Usted no está obligado a ser altruista, franco, una buena persona. Se le permite beneficiarse de la asimetría de la información. Si usted no pudiera hacer esto, se perdería el incentivo para descubrir la información acerca de los valores reales. (Es un ejemplo de la paradoja económica tradicional de que el vicio privado puede ser una virtud pública.) Y si por causas ajenas a la suya la otra parte del contrato no puede cumplir sus obligaciones, usted no está obligado a abstenerse de hacer cumplir sus derechos, incluso si el resultado de su insensibilidad obliga a la otra parte a declararse en quiebra. Es posible que *desee* abstenerse, por interés propio, porque los acreedores no suelen terminar bien en la quiebra. Pero si la ley *demandara* la contención, las tasas de interés serían más altas porque los derechos de los acreedores disminuirían.

La mención previa de la indemnización en esta sección ofrece la ocasión oportuna para introducir el concepto de "enriquecimiento injusto", otro ejemplo del uso del lenguaje moral de parte del derecho que el análisis económico puede hacer más perspicaz. La indemnización es el remedio están-

dar para el "enriquecimiento injusto", que tiene un significado más amplio que el solo aprovechamiento de la violación de los derechos de otra persona por parte nuestra. Supongamos que usted recibe un cheque por correo y, a pesar de que está a su nombre, esto es claramente resultado de un error. Si usted cobra el cheque y se queda con las ganancias, se dice que se enriqueció injustamente, y el destinatario previsto del dinero puede demandarle la indemnización. La injusticia es simplemente la desproporción entre el error y las consecuencias si se le permite quedarse con el dinero. El inconveniente de la devolución para usted es menor en relación con los costos de no haber cometido un error. Las personas se cuidarían en exceso si no hubiera límites en el derecho de una persona a tomar ventaja del error de otra persona. Adviértase la analogía con los límites a los derechos de los descubridores (antes, III.12).

Lecturas recomendadas

Baird, Douglas G., *Reconstructing Contracts* (2013).

Bar-Gill, Oren, *Seduction by Contract: Law, Economics, and Psychology in Consumer Markets* (2012).

Contract Law and Economics (Gerrit De Geest [comp.], 2011).

Craswell, Richard, "Offer, Acceptance, and Efficient Reliance", 48 *Stan. L. Rev.*, 481 (1996).

Katz, Avery, "The Strategic Structure of Offer and Acceptance: Game Theory and the Law of Contract Formation", 89 *Mich. L. Rev*., 215 (1990).

Kull, Andrew, "Reconsidering Gratuitous Promises", 21 *J. Legal Stud.*, 39 (1992).

Law and Economics of Insurance (2 vols.; Daniel Schwarcz [comp.], 2012).

MacLeod, W. Bentley, "Reputations, Relationships, and Contract Enforcement", 45 *J. Econ. Lit.*, 595 (2007).

Posner, Eric A., "Economic Analysis of Contract Law After Three Decades: Success or Failure?", 112 *Yale L. J.*, 829 (2003).

Posner, Richard A., "Let Us Never Blame a Contract Breaker", 107 *Mich. L. Rev.*, 1349 (2009).

_____, "The Law and Economics of Contract Interpretation", 83 *Tex. L. Rev.*, 1581 (2005).

Rasmussen, Eric, e Ian Ayres, "Mutual and Unilateral Mistake in Contract Law", 22 *J. Legal Stud*., 309 (1993).

Rea Jr., Samuel A., "Efficiency Implications of Penalties and Liquidated Damages", 13 *J. Legal Stud*., 147 (1984).

Schwartz, Alan, y Robert E. Scott, "Contract Interpretation Redux", 119 *Yale L. J.*, 926 (2010).

Scott, Robert E., "A Relational Theory of Default Rules for Commercial Contracts", 19 *J. Legal Stud.*, 597 (1990).

Shavell, Steven, "Damage Measures for Breach of Contract", 11 *Bell J. Econ*., 466 (1980).
Sykes, Alan O., "'Bad Faith' Breach of Contract by First-Party Insurers", 25 *J. Legal Stud.,* 405 (1996).
Vandenberghe, Ann-Sophie, "Behavioral Approaches to Contract Law", en *Contract Law and Economics*, 401 (Gerrit De Geest [comp.], 2011).

Problemas

1. Supongamos que los tribunales, al determinar los derechos y obligaciones de las partes de los contratos, no utilizan el criterio de la eficiencia para guiar sus decisiones, sino ciertos criterios no económicos de justicia. ¿Cuál sería el efecto de sus decisiones sobre el proceso de intercambio?

2. Paradine *vs*. Jane, Alleyn 26, 82 Eng. Rep. 897 (K. B. 1647), fue una acción para el pago de renta ejercitada por un terrateniente. La defensa del inquilino fue que había sido despojado por el ejército del príncipe Ruperto, un invasor. En términos económicos, ¿debiera el tribunal haber aceptado esta defensa?

3. Discuta las proposiciones siguientes:

a) Los daños de la confianza, antes que los de la esperanza, son la medida apropiada de los daños por el incumplimiento del contrato en los casos (por ejemplo, el error) en que no se llegó a un acuerdo.
b) La pérdida de la confianza corresponde en general al costo variable a corto plazo, mientras que la pérdida de la esperanza corresponde al costo variable a largo plazo (*cf.* después X.7).

4. ¿Por qué no es racional no pagar jamás nuestras deudas, y dejar que el acreedor incurra en el costo de los procedimientos de cobro, según la teoría de que el acreedor transigirá con la deuda en lugar de incurrir en tales costos? Véase Arthur Allen Leff, "Injury, Ignorance, and Spite-The Dynamics of Coercive Collection", 80 *Yale L. J.,* 1 (1970).

5. Si el aseguramiento es una función importante de los contratos, ¿cuáles remedios contractuales realizan esta función mejor (y peor), y debiera tener la parte ganadora el derecho de que la parte perdedora pagara sus honorarios legales?

6. A paga 100 dólares por un boleto del Super Bowl. Debido a cierta confusión provocada por el vendedor del boleto, éste no se entrega nunca a A, quien se pierde el juego. A habría pagado 10 000 dólares por el boleto. Supongamos que la confusión no era evitable por el vendedor a un costo razonable, pero que ha roto su contrato. ¿A cuánto debieran ascender los daños de A? Véase Samuel A. Rea Jr., "Non-Pecuniary Loss and Breach of Contract", 11 *J. Legal Stud*., 35 (1982).

7. La sección 3 de la Ley Harter, 46 U.S. C. 192, establece:

> Si el dueño de cualquier barco que transporte mercancías o propiedades desde o hasta cualquier puerto de los Estados Unidos de América ejerce la diligencia debida para volver a dicho barco en todos sentidos confiable y apropiadamente manejado, equipado y abastecido, ni el barco, ni su dueño o dueños, agentes o fletadores, se volverán o serán responsables de los daños o pérdidas resultantes de las faltas o errores de navegación o administración de dicho barco.

¿Puede imaginar alguna función económica posible de esta ley?

8. La ilegalidad (por ejemplo, una violación de la ley antimonopólica) es más probablemente una defensa contra el cumplimiento forzoso de un contrato ejecutorio que contra el cumplimiento forzoso de un contrato parcialmente ejecutado. ¿Por qué la diferencia?

9. ¿Es circular el requerimiento de la confianza *razonable* en los casos de impedimento colateral por promesa, según la teoría de que lo razonable de la confianza depende de la capacidad de cumplimiento forzoso de la promesa?

10. Evalúe el argumento siguiente: La Ley de Fraudes, que requiere la celebración por escrito de ciertos contratos y los declara de otro modo imposibles de cumplir por la fuerza, es innecesaria y paternalista porque cualquiera de las partes puede exigir, como una condición de la celebración del contrato, que se haga por escrito. Véanse Jason Scott Johnston, "The Statute of Frauds and Business Norms: A Testable Game-Theoretic Model", 144 *U. Pa. L. Rev.,* 1859 (1996); Eric A. Posner, "Norms, Formalities, and the Statute of Frauds: A Comment", 144 *U. Pa. L. Rev*., 1971 (1996).

11. ¿Debieran los contratos celebrados entre el gobierno y partes privadas regirse por las mismas reglas que los contratos plenamente privados? Vimos antes, en III.6, que la doctrina del uso público limita el poder del gobierno para celebrar contratos para la venta de propiedades públicas. Ninguna doctrina similar limita el poder de contratación de los dueños privados de propiedades. ¿Hay alguna razón económica para esta diferencia? ¿O para la regla de que, mientras que la autoridad aparente de un agente puede vincular a su principal en un contrato privado negociado por el agente, se requiere la autoridad efectiva para que un agente gubernamental sea capaz de vincular al gobierno con un contrato negociado por el agente? ¿Y es el argumento para la prohibición de las cláusulas punitivas en los contratos más fuerte o más débil en el caso de los contratos gubernamentales? Véase Daniel R. Fischel y Alan O. Sykes, "Governmental Liability for Breach of Contract", 1 *Am. Law & Econ. Rev.,* 313 (1999). Véase también Abraham L. Wickelgren, "Damages for Breach of Contract: Should the Government Get Special Treatment?", 17 *J. Law, Econ. & Organization,* 121 (2001).

12. Supongamos que A hace una gran inversión en la obtención de un contrato con B. Más tarde, B rompe el contrato y A demanda. ¿Debiera estar

facultado A para reclamar el pago de daños por la pérdida de esa inversión (confianza precontractual)? Véanse Lucian Arye Bebchuk y Omri Ben-Shahar, "Precontractual Reliance", 30 *J. Legal Stud*., 423 (2001); Richard Craswell, "Precontractual Investigation as an Optimal Precaution Problem", 17 *J. Legal Stud*., 401 (1988).

13. A y B celebran un contrato por escrito. Después, A demanda a B argumentando que durante las negociaciones B tergiversó de manera deliberada los beneficios que A obtendría del contrato. A no demanda por el incumplimiento de contrato, no puede. La regla de la prueba verbal prohibiría la declaración de que las promesas que se hicieron durante las negociaciones pero que no se reiteraron en el contrato deben considerarse contractualmente vinculantes. De modo que A demanda a B por agravio, acusándolo de fraude. La regla de la prueba verbal no es una regla del derecho de los cuasidelitos; pero supongamos que el contrato escrito contenía una cláusula que establecía que ninguna de las partes se basaba en declaración alguna que no se enunciara en el contrato escrito. Si es válida, la cláusula de "no confianza" (a menudo llamada cláusula del "niño mayor", ¿puede explicar por qué?) frustra la demanda de A por fraude porque no se puede obtener el pago de daños a menos que se basara en las declaraciones fraudulentas, y A ha renunciado a tal confianza. Así, aunque se da por sentado que B ha actuado indebidamente, A no tiene remedio, ya sea por contrato o por agravio de cuasidelito, si la cláusula del "niño mayor" es exigible. ¿Debiera serlo? Véase Extra Equipamentos e Exportação Ltda. *vs*. Case Corp., 361 F.3d, 359 (7th Cir., 2004).

14. En la discusión de los contratos de seguros, se dijo que "si una compañía de seguros no puede atribuir una probabilidad a un riesgo, no puede calcular la prima adecuada que se cobrará por asumir el riesgo". Y, sin embargo, el colectivo de seguros conocido como Lloyd's de Londres es famoso por asegurar contra la incertidumbre genuina (es decir, el riesgo que no se puede cuantificar), por ejemplo, el riesgo de que un concertista de piano sufra una lesión en las manos que termine con su carrera de manera prematura. Hay muy pocos concertistas de piano y demasiadas diferencias entre ellos (por ejemplo, en el cuidado que tienen para evitar un accidente) para hacer posible la cuantificación de tal riesgo. ¿Cuál es entonces el incentivo económico para redactar dicho seguro?

15. ¿Qué límites debe haber sobre la libertad de contratación? Considere lo siguiente: Cada vez hay más y más trasplantes de órganos de personas muertas (por ejemplo, de corazón, riñón, pulmones, córneas) a personas vivas. Desde un punto de vista económico, ¿sería una buena idea permitir que la gente celebre contratos vinculantes para la venta de sus órganos después de su muerte? Si es así, ¿qué normativas especiales podrían ser apropiadas que en un mercado más convencional serían innecesarias? Véanse Henry Hansmann, "The Economics and Ethics of Markets for Human Organs", 14

J. Health Politics, Policy & Law, 57 (1989); Lloyd R. Cohen, "Increasing the Supply of Transplant Organs: The Virtues of a Futures Market", 58 *Geo. Wash. L. Rev.*, 1 (1989); Michael Trebilcock *et al.*, "Testing the Limits of Freedom of Contract: The Commercialization of Reproductive Materials and Services", 32 *Osgoode L. J.*, 613 (1994). Y véase la discusión de la venta de bebés y la maternidad subrogada en el capítulo siguiente, V.5.

V. EL DERECHO FAMILIAR Y EL DERECHO SEXUAL[1]

La institución central de la familia es el matrimonio, una relación que se desenvuelve inciertamente en los límites del contrato. Por lo tanto, un capítulo sobre el derecho familiar es una continuación natural de uno sobre el derecho contractual, y el papel del derecho matrimonial en la canalización de la actividad sexual y procreativa convierte al análisis económico de la regulación de la actividad sexual en una extensión natural del análisis económico de la familia. Sin embargo, diversos temas de esta área general —la violación, la herencia (que es en gran medida familiar) y la discriminación sexual (un tema inseparable de la familia)— se pospondrán hasta capítulos posteriores.

V.1. La teoría de la producción doméstica

El análisis económico de la familia se funda en la percepción de que el hogar no es sólo una unidad consumidora, sino sobre todo una unidad productora de la sociedad. El alimento, la ropa, los muebles, las medicinas y otros bienes del mercado que compra la unidad familiar son insumos de la producción de nutrimento, calor, afecto, hijos y los otros bienes tangibles e intangibles que constituyen la producción de la unidad familiar. Un insumo fundamental de este proceso productivo no es un bien del mercado en absoluto; es el tiempo de los miembros de la familia, en particular el de la esposa (en la familia tradicional). (Con la aceptación cada vez mayor del matrimonio homosexual, el análisis ya no se puede limitar al matrimonio heterosexual; pero posponemos la discusión del matrimonio homosexual a la última sección del capítulo.)

La persistencia de la familia como una institución social le sugiere a un economista que la institución debe tener importantes propiedades de economización. ¿Cuáles podrían ser tales propiedades? ¿Las economías de escala (por ejemplo, compartir una cocina)? Pero tales economías podrían obtenerse, y a menudo se obtienen, fuera del matrimonio y en todo caso con frecuencia son menores que los costos de tener que adaptar nuestros gustos, horarios, etc., a los de otra persona. Un factor más importante es la división del trabajo, que genera ganancias derivadas de la especialización. En la familia tradicional el esposo se especializa en algún empleo del mercado que produce el ingreso que podrá emplearse en la compra de los bienes del mer-

[1] Véanse Judith Areen, *Cases and Materials on Family Law* (6ª ed., 2012); Homer H. Clark Jr. y Anne L. Estin, *Domestic Relations: Cases and Problems* (7ª ed., 2005); Carl E. Schneider y Margaret F. Brinig, *An Invitation to Family Law: Process, Problems and Possibilities* (3ª edición, 2006).

cado necesarios como insumos en la producción final de la unidad familiar, mientras que la esposa dedica su tiempo, en parte, al procesamiento de los bienes del mercado (comestibles, por ejemplo) para la producción de la unidad familiar (comidas, por ejemplo) y en parte a la producción de servicios fuera del mercado, como la crianza de los hijos. Al especializarse en la producción para el mercado, el esposo maximiza el ingreso monetario de la familia con el que se comprarán los bienes del mercado que necesita. Al especializarse en la producción doméstica, la esposa maximiza el valor de su tiempo como un insumo de la producción total de la unidad familiar. La división del trabajo tiende a maximizar el ingreso real total de la familia al permitir que el esposo y la esposa se especialicen en actividades complementarias. Es el mismo principio que nos lleva a esperar que una persona que trabaja medio tiempo como médico y medio tiempo como abogado produzca menos que la mitad de la producción total de servicios médicos y jurídicos de dos personas de habilidad igual a la suya, una de las cuales es un médico de tiempo completo y la otra un abogado de tiempo completo.

Las ganancias de la especialización se realizan a través del intercambio. El productor especializado no produce ni cerca de la gama completa de productos y servicios que le gustaría consumir; por lo que debe intercambiar su producción especializada con otros productores a fin de obtener esos bienes y servicios. Esto se advierte muy claramente en el ambiente matrimonial concentrándonos en el principal "bien" que produce el matrimonio: hijos. Muchos matrimonios no tienen hijos, pero pocos de ellos (en especial entre los primeros matrimonios) lo han decidido así; y es difícil creer que el matrimonio sería una institución común si la mayoría de la gente no quisiera tener hijos. La crianza de los hijos requiere, sobre todo en sus primeros años, una cantidad enorme de tiempo de los padres (tradicionalmente el de la madre), y una mujer que está ocupada en la crianza de un hijo no tendrá tiempo para trabajar en el mercado a fin de ganar el dinero que necesita para insumos complementarios (alimento, ropa, vivienda, etc.). Así, ella "intercambia" su trabajo en el hogar por el trabajo del esposo en el mercado; él "compra" el cuidado, por parte de ella, de los hijos que son de ambos.

La razón por la que es probable que el productor primario en el mercado sea el esposo y el productor primario en el hogar sea la esposa tiene dos caras. Primera, en la medida en que la discriminación sexual en el mercado laboral deprime los sueldos de las mujeres en relación con los de los hombres, el trabajo hogareño se vuelve más atractivo para la esposa que para el esposo. Segunda, en la medida en que las mujeres crían a los hijos y están mejor adaptadas, probablemente por razones de la biología evolutiva, al cuidado de los niños (especialmente los infantes), los costos de la producción hogareña serán menores para la esposa que para el esposo. Las mujeres que abandonan el trabajo para tener hijos y cuidarlos como infantes, aunque lo hagan porque esto es eficiente, tendrán salarios medios menores, aun cuando no haya dis-

criminación sexual, cuando regresen al mercado de trabajo. Es posible que sus habilidades en el mercado laboral pudieran haberse deteriorado por falta de uso, y al no haber esperado trabajar continuamente habrían invertido menos en su capital humano de mercado: su capacidad de obtener ingresos. El salario más bajo que pueden esperar como resultado de esta inversión menor reducirá el número de mujeres casadas que vuelven al mercado laboral. A medida que la participación en la fuerza laboral y los salarios de las mujeres han aumentado, los esposos han aumentado su participación en la producción de la familia, aunque todavía hay muy pocos esposos que se quedan en casa, además de los hombres desempleados cuyas esposas tienen empleo.

¡Qué anticuado nos ha llegado a parecer el modelo económico tradicional de la familia esbozado previamente! Veremos que para las mujeres se ha vuelto importante ir a la segura con sus apuestas sobre el matrimonio estableciéndose en el mercado laboral, a pesar del sacrificio que resulta de las ganancias de la especialización por parte de ambos cónyuges. Uno de los factores que reducen la tasa de matrimonios y aumentan la tasa de divorcios ha sido la disminución resultante en las ganancias del matrimonio. El matrimonio está dando paso a la cohabitación cada vez más; y, especialmente entre los afroamericanos y los hispanos, las familias monoparentales se están convirtiendo en la norma (la cabeza de la familia es la madre soltera). Más adelante discutiremos la disminución del matrimonio tradicional.

Si bien el matrimonio tradicional es una "sociedad", que se asemeja a una empresa pequeña, hay diferencias importantes entre la organización comercial y la familiar. Por ejemplo, la división del ingreso marital podría no estar determinada por el valor relativo de la contribución de cada cónyuge, como ocurriría en una sociedad mercantil; debido a la relación altruista entre los esposos, cada uno puede derivar una utilidad de los gastos de consumo de la otra. Una observación relacionada es que las tareas específicas dentro del hogar conyugal moderno occidental no están dirigidas y monitoreadas en una forma jerárquica, burocrática, o aun contractual. La razón es, de nuevo, el altruismo, un sustituto común en el matrimonio de los mecanismos de asignación y control en una empresa comercial. Altruismo es la condición en la que el bienestar de una persona es resultado del bienestar de otra. Si H ama a M (interpretando que el amor implica altruismo hacia la persona amada), un aumento de la felicidad o la utilidad de M será experimentado por H como un aumento de su propia felicidad o utilidad. Un punto adicional es que, en un hogar, la mayor parte del consumo es conjunto.

Aunque es un sustituto eficiente de la contratación formal, el altruismo no es el solvente para todo conflicto en la relación matrimonial.[2] Aunque H valúa la utilidad de M, es posible que no la valúe tanto como a su propia utilidad. Supongamos que el costo para M de hacer una hora más de trabajo

[2] Como se destaca en Gillian K. Hadfield, "Households at Work: Beyond Labor Market Policies to Remedy the Gender Gap", 82 *Geo. L. J.*, 89, 96-98 (1993).

hogareño es de 10, y que, según su percepción, el beneficio es de 8. Y supongamos que la utilidad de M interviene en la de H con un factor de descuento de 50%, de modo que el costo de esa hora extra para él es de sólo 5, y supongamos que su valor para él es de 9. Entonces, H tiene una ganancia neta de 4 derivada de la hora extra de su trabajo, y M una pérdida neta de 2. En el mercado esperaríamos que H compensara a M pagándole entre 2 (el precio mínimo que demandaría ella) y 4 (el precio máximo que pagaría H). Hay cierto apoyo empírico para un modelo del matrimonio basado en la negociación por oposición a un modelo en el que el jefe de familia perfectamente altruista asigna los recursos entre los miembros de la familia de acuerdo con las necesidades de cada uno de ellos. La experiencia indica que el consumo de las mujeres casadas, en relación con el consumo de los hombres casados, es mayor cuanto mayor sea la razón de los sexos (hombres a mujeres) y cuanto más favorables sean las leyes del divorcio para las mujeres; en otras palabras, cuanto mayor sea el poder de negociación de las mujeres casadas.[3]

La tasa de matrimonios ha disminuido considerablemente en los Estados Unidos y otras naciones occidentales. En 1950, el número de matrimonios en los Estados Unidos por cada mil habitantes era de 11; ahora es de 6.8, y es menor en los países de Europa Occidental. La tasa estadunidense de divorcios por cada mil habitantes se ha elevado a 3.6; sólo dos tercios de los matrimonios estadunidenses duran más de 10 años. Para muchos estadunidenses, la cohabitación ha tomado el lugar del matrimonio; 11% de las mujeres están cohabitando y 40% de los niños nacen fuera del matrimonio. Estas estadísticas sugieren que los beneficios netos del matrimonio para hombres y mujeres han disminuido. El valor de tener más de uno o dos hijos ha declinado con la abolición del trabajo infantil, la disminución del número de familias de agricultores (los padres pueden emplear productivamente a sus hijos en las granjas sin violar las leyes de trabajo infantil), la sustitución del apoyo que les prestan a las personas mayores sus hijos por la seguridad social y la disminución de la mortalidad infantil (se requieren menos nacimientos para tener una confianza razonable en la obtención de nuestro número deseado de hijos adultos). Pero estos cambios no son nuevos. Lo nuevo es que el costo de los hijos ha aumentado notablemente, en especial el costo de oportunidad para la madre, debido a que el advenimiento de aparatos hogareños baratos que ahorran trabajo y la creciente fracción de los empleos que no requieren gran fuerza física han disminuido para las mujeres el costo de trabajar en el mercado y aumentado la demanda de sus servicios por parte de los empleadores. El consiguiente aumento del ingreso neto que pueden ganar las mujeres en el mercado (salarios menos costos del trabajo, incluidos los costos del tiempo que se dedica a la producción hogareña) ha elevado el costo de oportunidad del hecho de ser un ama de casa, porque ese costo es el ingreso neto

[3] Pierre-André Chiappori, Bernard Fortin y Guy Lacroix, "Marriage Market, Divorce Legislation, and Household Labor Supply", 110 *J. Pol. Econ.*, 37 (2002).

de mercado sacrificado por tener que permanecer en el hogar. Hoy en día, cerca de 54% de las mujeres adultas tienen un empleo, en comparación con 68% de los hombres; y el ingreso promedio de una mujer con un empleo de tiempo completo constituye 78% del ingreso promedio de un hombre con un empleo de tiempo completo.

Un aumento en el costo de oportunidad de la producción doméstica de una mujer aumenta el precio sombra de los hijos en la familia. Además, al ser más capaces para mantenerse a sí mismas que en el pasado, las mujeres son más selectivas respecto a la búsqueda de pareja, y esto ha reducido la tasa de casamientos. Y existe un mayor acceso a relaciones sexuales antes del matrimonio; casarse *para* tener sexo se está convirtiendo en cosa del pasado a medida que desaparecen los prejuicios. Los cambios en la legislación, por otra parte, han reducido la diferencia entre el matrimonio y los sustitutos cercanos, como la cohabitación: los hijos que nacen fuera del matrimonio ya no están sometidos a las desventajas asociadas con la "ilegitimidad", como la negación de derechos a la herencia, y el divorcio incausado ha reducido su costo. Todo esto hace de la cohabitación un sustituto más cercano del matrimonio, con la ventaja de permitir "asegurarse contra la incertidumbre relativa a la calidad de la pareja en la relación".[4] Entre más escasas sean las ganancias derivadas del matrimonio, más preponderante será la "calidad de la pareja" como un factor en la relación, sobre todo porque, cuando ambos cónyuges trabajan en el mercado, las ganancias de un integrante de la pareja que se especializa en la producción del mercado y las del otro que se especializa en la producción doméstica se reducen; de este modo, la relación es menos productiva y, por ende, es menos probable que continúe si el afecto mutuo disminuye.

Tendemos a tratar a los hijos como un "bien" final, pero es posible tratarlos como un insumo de otros bienes. El economista especula que los hijos se producen: *1)* como un subproducto no buscado de la actividad sexual, *2)* como una inversión generadora de ingreso, *3)* como una fuente de otros servicios para los padres, y *4)* (realmente un subconjunto de *3*) como resultado de un instinto o deseo de preservar a la especie o de perpetuar las características genéticas, el apellido o la memoria de los padres, o *5)* porque simplemente es lo que se espera. En una época de amplia disponibilidad de anticonceptivos y abortivos, *1)* se ha vuelto relativamente poco importante; nunca fue *muy* importante, excepto allí donde la ley o la costumbre confinaban la actividad sexual al matrimonio. *2)* fue alguna vez importante en nuestra sociedad (como lo es ahora en las sociedades muy pobres) pero, como se dijo antes, ya no lo es. Esto ha vuelto obsoleto a *2)* y desatado una búsqueda de servicios menos tangibles que los padres podrían obtener de un hijo (el respeto, por ejemplo). *3)* a *5)* son las explicaciones más plausibles del deseo de tener hijos

[4] Ahu Gemici y Steve Laufer, "Marriage and Cohabitation", 2012, economics.sas.upenn.edu/system/files/event_papers/cohabitation_031212.pdf?

en la sociedad contemporánea, pero *5)* disminuye a medida que las tasas de matrimonios y de natalidad descienden. El gusto por los hijos es un subconjunto de *3)*: el placer que obtenemos de la presencia de nuestros hijos y de nuestros nietos es el resultado del "consumo" de los "servicios" intangibles que ellos nos brindan.

V.2. Formación y disolución del matrimonio

Las sociedades mercantiles son asociaciones contractuales voluntarias, al igual que los matrimonios hasta cierto punto. El "mercado de matrimonios" es una metáfora adecuada para el elaborado proceso de búsqueda mediante la cual los individuos tratan de encontrar parejas maritales con quienes formar hogares productivos. El mercado matrimonial es racional. Por ejemplo, los hombres brillantes tienden a casarse con mujeres brillantes; una analogía agrícola revelará por qué. Supongamos que hay dos granjas, y el suelo de una de ellas es (sin tratamiento) dos veces más productivo que el suelo de la otra (sin tratamiento). Un producto químico duplicará el rendimiento de cualquiera de las granjas a la que se aplique, pero sólo hay suficiente para una granja. ¿Debiera aplicarse a la granja de suelo más pobre, según la teoría de que esa granja lo necesita más? ¿Debiera dividirse entre las granjas? (La mitad del químico aumentará la producción de una granja en 50%.) ¿O debiera aplicarse todo a la granja de suelo más rico? Esto último. Supongamos que la producción de la granja con el suelo más rico es (antes del tratamiento con el químico) 2, mientras que la producción de la otra granja es 1. Si se aplica el químico en su totalidad a la granja más pobre, la producción total de las dos granjas será 4 (2 + 2); si se aplica por partes iguales, será 4.5 (3 + 1.5); pero si se aplica en su totalidad a la granja mejor, será 5 (4 + 1). Asimismo, esperaríamos que los despachos de abogados que tienen los socios mejores contrataran a los mejores asociados para trabajar con ellos, que los mejores jueces tuvieran a los mejores asistentes jurídicos, que las escuelas de derecho con los mejores estudiantes tuvieran a los mejores profesores y que las empresas mercantiles de los mercados prósperos tuvieran en promedio mejores ejecutivos que las empresas de los mercados declinantes. Y lo mismo ocurre en el matrimonio si las cualidades positivas de los cónyuges se relacionan multiplicativamente como en los ejemplos de la granja, la empresa legal, los asistentes jurídicos, la escuela de derecho y la corporación, y no sólo aditivamente.

¿O la analogía apropiada para el matrimonio debiera ser, por el contrario, el comercio internacional de bienes disímiles —digamos, trigo por aviones—, en el que no hay ninguna presunción de que lo mejor del uno se está intercambiando por lo mejor del otro, lo segundo mejor del uno por lo segundo mejor del otro, etc.? Estos bienes no se usan juntos, como en los ejem-

plos anteriores; por lo tanto, no existe el efecto multiplicador potencial que en el ejemplo del matrimonio podría incluir la producción de hijos de inteligencia o belleza superior. Hay otra razón para la reunión "de coordinación positiva" de los seres humanos: favorece la reducción de la fricción y, por ende, de los costos de transacción dentro del hogar. Sin embargo, la compensación de este hecho es la posibilidad reducida de especialización cuando los cónyuges tienen preferencias muy similares.

A pesar de la semejanza del matrimonio con una sociedad mercantil, la relación marital no es un ejemplo puro de los principios del libre mercado, o por lo menos no lo era antes del divorcio incausado (del que nos ocuparemos más adelante). Tres características en particular, que parecen extrañas desde la perspectiva del capítulo anterior e incompatibles entre sí, distinguen al derecho matrimonial tradicional del derecho contractual.

Primero, las partes no están en libertad para fijar la duración del contrato o siquiera para terminar el contrato por mutuo consentimiento; la duración es de por vida, sujeta a la terminación por una causal legal. Hoy en día, cualquiera de las partes puede divorciarse por mutuo acuerdo o consentimiento —eso es lo que significa el divorcio incausado—, pero estamos describiendo el derecho matrimonial tradicional. Y los matrimonios con una duración establecida con antelación, digamos cinco años, aún están prohibidos.[5]

Segundo, la sanción por el incumplimiento es más severa que en el caso de un contrato regular. Si el esposo abandona a la esposa, no sólo deberá continuar manteniéndola (lo que es análogo a tener que pagar daños y perjuicios por incumplimiento de contrato), sino que no podrá casarse con nadie más. Esto es como si pudiera prohibirse al violador de un contrato la celebración de un contrato sustituto del que había roto… y por el resto de su vida. En realidad, los incumplimientos de contratos de servicios personales no pueden prohibirse en absoluto, a menos que los servicios del promitente sean únicos y esto volviera inadecuado el remedio del pago de daños para el promisario. A pesar de la naturaleza fuertemente vinculada de su relación, si los cónyuges tienen una disputa durante el curso del matrimonio los tribunales no intervendrán de ordinario para arreglar la disputa, aunque los cónyuges hayan firmado un contrato que asigne derechos y obligaciones que uno de los cónyuges alegue que han sido violados; los cónyuges estarán obligados a resolver su disputa por sí mismos o a separarse.

En consecuencia, el derecho matrimonial tradicional es una amalgama desconcertante de intrusión legal (en lo tocante a la duración del contrato y a las sanciones por el incumplimiento) y de no intervención legal. La explicación podría ser que el "contrato" matrimonial afecta a terceros que no

[5] En cambio, a los musulmanes chiitas se les permite celebrar matrimonios temporales por un periodo determinado, pero este periodo no debe ser menor a tres meses; esto para facilitar la determinación de la paternidad.

han consentido: los hijos del matrimonio. Por supuesto, incluso en un sistema de divorcio consensual, los padres que aman a sus hijos tomarán en cuenta el costo del divorcio para los hijos al decidir si se divorcian; pero, a menos que sean completamente altruistas, no interiorizarán plenamente el costo para los hijos, de modo que podrían divorciarse aunque los costos totales para todos los involucrados superaran a los beneficios totales. La educación pública, las leyes de trabajo infantil, las leyes contra el abuso infantil y las obligaciones de apoyo son ejemplos de políticas que reconocen y buscan corregir deficiencias graves del altruismo de los padres hacia los hijos.

Podría parecer que, si se encierra a los padres en un matrimonio miserable, volviendo imposible el divorcio, se condenará también a los hijos a la miseria. Pero esto olvida el hecho de que la prohibición del divorcio inducirá una búsqueda más cuidadosa de una pareja matrimonial en primer lugar. Entre más costoso sea un error, menos probable será que se cometa; y un error en la elección de un cónyuge es más costoso en un sistema que prohíbe el divorcio (o lo vuelve muy difícil o costoso) que en un sistema que lo permita. Además, entre más larga sea la búsqueda de pareja, mayor será la edad media de los cónyuges al contraer matrimonio; y las personas más maduras y experimentadas tienen menos probabilidades de cometer errores que las novatas. Tenemos así la curiosa paradoja de que, al volver más difícil el divorcio, ¡podrían promoverse, en efecto, matrimonios felices!

Mientras más fácil sea divorciarse, menos se diferenciará el estado matrimonial de la soltería; por lo tanto, mientras más bajo sea el costo de posponer el matrimonio, será más bajo el costo de una búsqueda prematrimonial prolongada, y la búsqueda más prolongada podría generar una *reducción* en la tasa de divorcios de las parejas casadas luego de que el divorcio incausado remplazara al divorcio por causal.[6] De acuerdo con esta sugerencia, un estudio revela que las leyes del divorcio incausado generan un aumento en la tasa de divorcios de las parejas que se casaron antes de que se adoptaran las leyes, pero que las parejas que se casaron después de que se aprobaran las leyes se divorcian menos porque una probabilidad mayor de divorcio genera una búsqueda matrimonial más cuidadosa.[7] Aun así, es difícil creer que la búsqueda prematrimonial prolongada inducida por el régimen incausado produzca matrimonios tan estables como aquellos que requieren el compromiso que entraña el matrimonio cuando es difícil divorciarse.

Otra diferencia entre el divorcio por causal y el divorcio incausado es que este último proporciona la opción de salir rápidamente, en especial valiosa para las esposas. Un estudio empírico reveló que la aprobación de las leyes

[6] Imran Rasul, "Marriage Markets and Divorce Laws", 22 *J. Law, Econ. & Organization*, 30 (2005).

[7] Stéphane Mcchoulan, "Divorce Laws and the Structure of the American Family", 35 *J. Legal Stud.*, 143 (2006).

del incausado redujo la violencia de los esposos hacia las esposas, así como la tasa de suicidios de las esposas.[8]

El hecho de que la probabilidad de que la búsqueda matrimonial se prolongue bajo casi cualquier régimen de divorcio podría explicar la renuencia de la ley bajo el régimen basado en la causa —tan diferente del derecho contractual— a reconocer el fraude como una razón para anular un matrimonio, a menos que el fraude involucre al sexo (típicamente, el esposo no reveló antes del matrimonio que era impotente). La regla refleja y fomenta un cortejo prolongado, lo que provee una oportunidad para que ambos cónyuges descubran los disfraces con que tratan de hacerse aparecer en las relaciones personales como poseedores de mejores cualidades y perspectivas que las que realmente tienen. Entre mayor sea la búsqueda precontractual, menores serán los remedios legales necesarios. Pero el fraude en relación con el sexo no sólo afecta al meollo del contrato matrimonial; era difícil de detectar cuando las relaciones sexuales prematrimoniales eran mal vistas, y la disolución de un matrimonio sin hijos implica costos sociales mínimos.

Hasta bien entrado el siglo XIX, la negación del derecho inglés a reconocer *cualquier* razón para el divorcio podría haber protegido al cónyuge más débil (invariablemente la esposa) con mayor eficacia que si se hubiese permitido el divorcio por causal. En un sistema que permite este régimen, un esposo que desea "salirse" del matrimonio tendrá un incentivo para tratar tan mal a su esposa que ella se vea obligada a demandar el divorcio, siempre que el pago de pensión alimenticia u otros remedios no arrojen los costos totales del maltrato sobre los hombros de él, como no ocurría a menudo en un sistema en el que el litigio era lento, costoso e incierto. Pero si pueden superarse las dificultades remediales (y, por lo tanto, se reduce la importancia de la opción de salida rápida que discutimos anteriormente), el establecimiento del divorcio por causal tiene sentido económico, ya que permite por lo menos una comparación burda entre los costos del divorcio para los hijos y el costo de permanecer casado para un cónyuge severamente maltratado. Por otra parte, las causas tradicionales del divorcio, con una excepción parcial y muy interesante (el adulterio), parecen haberse limitado a los casos en que la mala conducta del esposo tendía a dañar a los hijos tanto como a la esposa: los casos de la enfermedad mental, la crueldad extrema y la criminalidad.

Tradicionalmente, un solo acto de adulterio de la esposa era causal de divorcio, mientras que el esposo tenía que ser un adúltero habitual para que la esposa tuviera derecho al divorcio. Una explicación económica es que el adulterio de la esposa es más costoso para el esposo que el caso inverso. Una esposa adúltera podría concebir un hijo que no es de su esposo; y, dado que la capacidad de las mujeres para criar hijos es claramente finita, los beneficios

[8] Betsey Stevenson y Justin Wolfers, "Bargaining in the Shadow of the Law: Divorce Laws and Family Distress", 121 *Q. J. Econ.*, 267 (2006).

del matrimonio para el esposo se verán significativamente perjudicados, si suponemos que desea tener hijos de su propia sangre. En cambio, el adulterio del esposo no reduce necesariamente el número de hijos que su esposa criará o el sostén que él otorgará a cada uno de ellos; los beneficios del matrimonio para ella, por lo menos en lo tocante a los hijos, podrían no verse perjudicados en absoluto (aunque podrían hacerlo si él engendra y sostiene a hijos ilegítimos). Pero si el esposo es un adúltero *habitual*, podría desatender las necesidades de su esposa y sus hijos (legítimos) (y, por lo tanto, probablemente engendrar a un hijo ilegítimo) hasta el punto de imponer a la esposa un costo tan elevado como el que impondría al esposo el adulterio aislado de una esposa.

El problema del reconocimiento de *cualquier* razón para el divorcio es que así se mina la política, plausiblemente diseñada para la protección de los hijos del matrimonio, contra las disoluciones voluntarias del matrimonio. Un acuerdo para disolver un matrimonio involucra sólo a dos personas; aunque hay un problema de monopolio bilateral, los costos de transacción no debieran ser prohibitivos. Una vez que las partes han llegado a términos mutuamente aceptables, sólo necesitan fabricar pruebas de una conducta que provea una base legal para el divorcio a fin de eludir una ley que prohíba el divorcio consensual. La fabricación de pruebas no carece de costos, de modo que una ley de divorcio exigente preservará algunos matrimonios al aumentar los costos de la disolución. Pero si la sociedad estuviera decidida a preservar los matrimonios, por lo menos impediría que las partes controlaran las pruebas; permitiría el divorcio sólo cuando el agente del ministerio público, o algún otro tercero, hubiesen probado que se ha incumplido el contrato matrimonial. Conforme han disminuido las ganancias del matrimonio, debido principalmente a mayores oportunidades de empleo para las mujeres en el mercado y a la reducción resultante en la demanda de hijos, ha aumentado la presión en favor del divorcio. Esto ha vuelto cada vez más costosa la aplicación de una política contra el divorcio consensual, lo que abrió camino hacia leyes de divorcio más liberales.

El divorcio incausado, que al permitir a cualquiera de los cónyuges terminar el matrimonio a voluntad es un paso más allá del divorcio de mutuo acuerdo, convierte al matrimonio en una relación muy similar a la de la asociación. ¿Debiera el derecho dar el paso siguiente y permitir que las partes del matrimonio especifiquen en el momento de contraerlo su duración, sus respectivos derechos y obligaciones, así como las razones de su terminación? De ser así, ¿debieran ser el único remedio para el incumplimiento la disolución del matrimonio, o debieran los tribunales otorgar pagos de daños como en el caso de un juicio ordinario de incumplimiento de contrato? Estas preguntas están relacionadas entre sí. Una objeción económica importante para convertir el matrimonio en un régimen contractual puro es la dificultad de determinar el pago de daños por incumplimiento, así como el hecho de que

la relación de monopolio bilateral entre los cónyuges debería facilitar la resolución informal de las disputas. Pero si el único remedio legal para el incumplimiento es la disolución del matrimonio se disminuyen las objeciones al matrimonio contractual; no se recurrirá a los tribunales para que valúen los daños cada vez que los cónyuges tengan una disputa sobre el cumplimiento del contrato.

Sin embargo, es menos probable que el divorcio consensuado tienda a ser equilibrado en comparación con el divorcio por causal o el divorcio incausado. El divorcio consensuado crea problemas de monopolio bilateral que el divorcio incausado previene. En el divorcio consensuado un cónyuge puede obligar al otro a hacer concesiones al negarse a aceptar el divorcio. En el divorcio incausado aún se da una negociación —la parte que no quiere el divorcio puede tratar de compensar a la otra para que permanezca en el matrimonio—, pero no hay ninguna posibilidad de retención. En el divorcio por causal no es necesaria la negociación, sólo una determinación judicial de la culpa.

¿Podrían las fuerzas discutidas en esta sección disminuir la tasa de matrimonios hasta llegar a cero, o casi a cero, sustituyendo al matrimonio por un régimen contractual puro: el triunfo de la libertad de contratación? Es improbable. El matrimonio continúa señalando el compromiso de una manera que ninguna otra relación adulta lo hace, aunque sólo sea porque las parejas que se casan exageran la probabilidad de no divorciarse. Poner fin a un matrimonio es social y emocionalmente mucho más difícil que terminar una cohabitación, en parte debido a las expectativas exageradas que la gente trae al matrimonio (un ejemplo de sesgo de optimismo, una deficiencia cognitiva común), en parte porque el divorcio se sigue considerando ampliamente como un signo de fracaso y en parte porque el divorcio seguido de nuevas nupcias complica la relación con los hijos del matrimonio anterior. Y es difícil imaginar sustitutos contractuales satisfactorios que crearan un compromiso comparable, lo que tendría que definir las obligaciones y su desempeño satisfactorio, y especificar las sanciones por incumplimiento. El compromiso con el matrimonio, como la prohibición del divorcio, fomenta la cooperación y las inversiones ligadas específicamente a la relación entre las partes.

V.3. Consecuencias de la disolución

Cuando se disuelve una asociación convencional, los activos de la asociación deben distribuirse entre los socios, y lo mismo ocurre con el matrimonio. Pero es difícil determinar las porciones respectivas de los activos adquiridos por la unidad familiar. Si la esposa ha tenido muy escaso ingreso del mercado, todos o la mayor parte de los activos tangibles de la unidad familiar han sido comprados con el dinero del esposo. Pero es posible que la ca-

pacidad de éste de ganar ingresos deba mucho a los esfuerzos de ella. Es posible que ella lo haya mantenido mientras él estudiaba en la escuela de derecho o de medicina, reduciendo su propio consumo y también sacrificando oportunidades para aumentar su propia capacidad de ganar ingresos mediante una capacitación avanzada. Habiendo invertido en la capacidad de ganar ingresos de él, ella tiene derecho a que se le pague con los activos o los ingresos de él, como los tribunales y las legislaturas lo están empezando a entender.[9] Pero ¿qué pasa si el divorcio ocurre muchos años después de la celebración del enlace: no podría haber sido retribuida la esposa en su totalidad durante el matrimonio? Sí, si la inversión que hizo en la capacidad de generar ingresos de su marido se concibe como análoga a un préstamo, antes que como una inversión de capital. Ésta no se "rembolsa"; el inversionista conserva su interés hasta que lo vende o hasta que la empresa lo liquida.

Supongamos que, debido a restricciones de liquidez y a la dificultad para obtener un préstamo con la garantía de un activo incierto, el esposo que se divorcia no puede pagar a la esposa una suma global igual al valor de la participación de ella en el título profesional de él. ¿Tendría más sentido tratarla a ella como a un inversionista en acciones y otorgarle, en consecuencia, un porcentaje de las ganancias de él a medida que ocurran, o tratarla a ella como un prestamista que tiene derecho a un pago periódico fijo? En el primer caso, el esposo podría tener, después del divorcio, un incentivo disminuido para trabajar duro porque él sólo se queda con una parte de lo que gana, por tiempo indefinido; pero en el segundo caso podría tener un incentivo para asumir en ese periodo riesgos financieros excesivos porque el prestamista comparte las pérdidas con el prestatario, pero no las ganancias.[10]

Aun si la pareja se casó después de que el hombre completara su capacitación profesional, y la esposa se especializó exclusivamente en la producción doméstica, sería un error atribuir todos los activos de la unidad familiar a la producción del esposo. El valor de la contribución no pecuniaria de la esposa podría haber sido tan grande como el de la contribución del esposo,[11]

[9] Véanse, por ejemplo, Harris y Harris, 244 P.3d 801, 810-811 (Or. 2010); Libby *vs*. Libby, 781 A.2d 773, 775 (Me. 2001); Holterman *vs*. Holterman, 762 N.Y.S.2d 152, 153-154 (App. Div. 2003); Fla. Stat. 61.075(1); Iowa Code 598.21(5) (c); y para su discusión, Daniel D. Polsby y Martin Zelder, "Risk-Adjusted Valuation of Professional Degrees in Divorce", 23 *J. Legal Stud.*, 273 (1994).

[10] Véase Katherine Wells Meighan, "For Better or for Worse: A Corporate Finance Approach to Valuing Educational Degrees at Divorce", 5 *George Mason L. Rev.,* 193 (1997). *Cf*. después con XV.5 los XIV.4 y XXI.9.

[11] Se ha estimado que, en promedio, el valor de la producción doméstica de la esposa es igual a más de 70% del ingreso monetario (es decir, del mercado) de la unidad familiar después de impuestos, lo que implica que la producción doméstica de la esposa genera 41% del ingreso total de la unidad familiar (0.7/1.7). Reuben Gronau, "Home Production: A Forgotten Industry", 62 *Rev. Econ. & Stat*, 408 (1980).

y en todo caso es improbable que haya sido insignificante. Si el esposo hubiera tenido que dedicar una parte sustancial de su tiempo a la producción doméstica, su ingreso de mercado habría sido menor y se habrían acumulado menos activos; algunos de tales activos se compraron entonces por el trabajo de la esposa en la unidad familiar.

Sin embargo, es arbitraria la regla establecida en los patrimonios de propiedad comunitaria en el sentido de que 50% de los activos acumulados durante el matrimonio se consideran propiedad de la esposa tras la disolución, sobre todo cuando se aplica al creciente número de individuos excesivamente ricos de los Estados Unidos. Supongamos que dos personas se casan cuando son muy jóvenes, y que no llevan al matrimonio ningún activo financiero. H se vuelve un empresario fabulosamente exitoso. Cuando llega a los 50 años de edad ha acumulado un patrimonio neto de 150 millones de dólares. M permanece en el hogar y cría a los hijos de la pareja, aunque con ayuda considerable de personal remunerado. Si H y M se divorcian ahora (supongamos que sin causal de ninguna de las partes), ¿a qué debiera tener derecho M? Podemos enfocar este interrogante preguntándonos cuánto habría tenido que pagar H para contratar a alguien (o a varias personas) para proveer los servicios que M proveyó (quizá no mucho, si M ha recibido ayuda considerable del personal al que H pagó). La interrogante puede precisarse preguntando cómo habrían convenido H y M dividirse sus activos (en caso de un divorcio) cuando se casaron, si hubiesen considerado la posibilidad de que H se volviera fabulosamente exitoso y contemplara esta contingencia en un acuerdo prenupcial. Si bien el éxito de H en los negocios podría deberse en buena parte a las contribuciones de M aunque ella nunca trabajara en el mercado, es improbable que la dimensión *fabulosa* de su éxito se debiera a tales contribuciones, lo que sería un argumento en contra de que ella obtuviera algo cercano a 75 millones de dólares. Tal derecho sería un pago excesivo para ella, y es de presumirse que H no habría estado de acuerdo con esto *ex ante*. Este enfoque contractual hipotético para la cuestión de la valuación podría volverse operativo examinando los términos de los acuerdos prenupciales para ver las disposiciones que contienen de ordinario respecto al éxito financiero inesperado de cualquiera de los cónyuges. Esto proveería un indicador del mercado para usarlo en los casos en que no hubo ningún acuerdo prenupcial.[12]

Se ha sugerido que un cónyuge que planee especializarse en la producción hogareña debiera recibir un gravamen legalmente exigible ("acuerdo de seguridad premarital") sobre el ingreso del cónyuge (más probablemente a favor de la esposa) en la medida en que es posible considerar en térmi-

[12] ¿Debiera reducirse la participación de la esposa en los activos maritales si la disolución del matrimonio se debiera a su propia culpa? En principio sí, y se reduciría por los costos que la esposa impuso al esposo por la disolución del matrimonio.

nos monetarios el valor de esa producción.[13] ¿Es más probable que las mujeres se vean favorecidas o perjudicadas por la mercantilización explícita de las relaciones maritales en esta forma? En general, no existe una objeción económica a la materialización, pero podemos pensar en instancias particulares en las que el mercado es una alternativa ineficiente al contrato, por ejemplo, la gobernanza de las relaciones entre los cónyuges, como mencionamos anteriormente.

Además de prescribir la división de los activos maritales, una sentencia de divorcio ordenará por lo común al esposo que pague a la exesposa: *1)* una suma fija periódica, a menos que y hasta que ella se case de nuevo (pago de pensión alimenticia), y *2)* una parte del costo de la crianza de los hijos del matrimonio, de quienes tendrá ella de ordinario la custodia (sostenimiento de los hijos).[14] El pago de pensión alimenticia al excónyuge es analíticamente complejo. Pareciera desempeñar tres funciones económicas distintas:

1. Es una forma del pago de indemnización por incumplimiento del contrato matrimonial. Pero, si esto fuera todo, sería de esperarse que el pago de pensión alimenticia se realizara en una sola exhibición, como otras formas de los daños y perjuicios, a fin de minimizar los costos de la supervisión judicial; y nunca se otorgaría al cónyuge culpable, como ocurre con frecuencia.
2. Es un método de compensar a la exesposa (en el matrimonio tradicional) por su participación en los activos de la sociedad conyugal. A menudo, el activo principal al que la esposa habrá contribuido por su trabajo en el hogar o en el mercado —como en nuestro ejemplo de la esposa que mantuvo a su esposo mientras él iba a la universidad a obtener un posgrado— es la capacidad de ganar ingresos del esposo. Éste es un activo contra el cual resulta difícil obtener préstamos; el prestamista no puede ejecutar la hipoteca esclavizando al prestatario. Y el activo contra el que se obtiene el préstamo —la capacidad del prestatario de generar ingresos— es inherentemente incierto y es probable que la compensación se tenga que retrasar por muchos años en caso de que un periodo largo de enseñanza o aprendizaje preceda a las ganancias sustanciales. Por lo tanto, podría resultar inviable que el esposo obtuviera el dinero necesario para comprarle a la esposa, en una suma de una sola exhibición, la parte del activo que ella puede

[13] Martha M. Ertman, "Commercializing Marriage: A Proposal for Valuing Women's Work through Premarital Security Agreements", 77 *Tex. L. Rev.*, 17 (1998). Véase también Katharine Silbaugh, "Commodification and Women's Household Labor", 9 *Yale J. Law & Feminism,* 81 (1997).

[14] Se ha vuelto común la custodia conjunta. Como las mujeres pasan más tiempo en el empleo del mercado, los hombres deben dedicar más tiempo a la producción doméstica; cada uno se vuelve menos especializado y, por lo tanto, la superioridad de la mujer en la custodia de los hijos se vuelve menos marcada.

reclamar en justicia como suyo en virtud de sus contribuciones; en lugar de ello, él deberá pagarle a ella tiempo extra con la corriente de ingresos generada por el activo. Pero ésta no es una explicación completa del pago de pensión alimenticia a la esposa; porque si así fuera la ley no terminaría con esos pagos cuando la esposa vuelve a casarse.

3. La función económica más importante del pago de pensión alimenticia podría ser la de proveer a la esposa de una forma de pago de indemnización por despido o de subsidio de desempleo. En la familia tradicional, en la que la esposa se especializa en la producción hogareña, se deprecian todas las habilidades que pueda haber tenido en la producción para el mercado, y finalmente sus posibilidades primarias de empleo —si se disolviera el matrimonio actual— se reducirían a la perspectiva de volver a casarse y formar una unidad familiar nueva en la que pudiera desempeñar su oficio. En virtud de que la búsqueda de un cónyuge adecuado es a menudo prolongada, y dado que la edad podría depreciar la capacidad de una persona para formar un nuevo matrimonio que le produzca tanto ingreso real como el anterior, tiene sentido que se incluya como un término convencional del contrato matrimonial implícito una forma del pago de indemnización por despido para que la esposa mantenga un nivel de vida durante la búsqueda de un nuevo esposo (la que puede ser inútil, en cuyo caso el pago de pensión alimenticia continuará por el resto de su vida). Consideremos una analogía con la práctica legal. Al convenir en trabajar para una firma legal que se especializa exclusivamente en la negociación de hipotecas de tanqueros petroleros, a un abogado podría dificultársele en extremo, en caso de que algún día fuera despedido, el encuentro de una posición igualmente remunerativa: sus habilidades en empleos jurídicos alternativos se habrían depreciado en virtud de su concentración en una especialidad limitada. Pero ésa es una razón especialmente poderosa para que demande, como una condición para trabajar en tal empresa, que se le asegure que en caso de ser despedido se le continuará pagando su sueldo hasta que encuentre un trabajo igualmente remunerativo, aunque la búsqueda sea prolongada.

Una alternativa en el caso del ama de casa y en el del abogado sería un salario mayor para compensar el riesgo de un desempleo prolongado en caso de despido. Pero el esposo podría estar imposibilitado para hacer a la esposa los pagos de transferencia necesarios, sobre todo durante los primeros años del matrimonio, cuando la unidad familiar podría no tener grandes activos líquidos. Además, dado que la fecha del divorcio y el periodo (posterior al divorcio) destinado a obtener un nuevo matrimonio no se podrían estimar con antelación, la compensación *ex ante* desplazaría el riesgo sustancial a la esposa divorciada. La incertidumbre de las segundas nupcias también es una

razón de peso para otorgar el pago de pensión alimenticia sobre una base periódica antes que como un pago global.

Así como el pago de indemnización por despido es generalmente independiente de que el empleador tuviera la culpa de despedir al trabajador, los pagos de pensión alimenticia, vistos como una forma del pago de indemnización por despido, no son dependientes de las nociones de una culpa. Pero, así como un empleado podría perder su derecho al pago de indemnización por despido por haber renunciado en incumplimiento de su contrato de empleo, el pago de pensión alimenticia debiera negarse o reducirse (y así ocurre a veces) si la esposa incurriera en grave culpa al pedir la disolución del matrimonio. Mejor aún: todos los daños que causara a la familia por separarse del matrimonio podrían restarse de sus pagos de pensión alimenticia si su participación en los activos maritales fuese insuficiente para cubrirlos.

¿Acaso la aparición del divorcio incausado volvió más o menos importante que los tribunales adoptaran un enfoque económico cuidadoso para el cálculo del pago de pensión alimenticia? Más. La razón es que, cuando el divorcio se basa en una causa, el hombre que quisiera el divorcio tendría que obtener el consentimiento de su esposa, y esto le permitiría a ella negociar un pago generoso de pensión alimenticia.

V.4. La protección legal de los hijos

A fin de realizar su potencial como adultos —en términos económicos, a fin de alcanzar un nivel elevado de utilidad de por vida—, los hijos requieren una inversión considerable de tiempo parental y de insumos del mercado (alimentos, ropa, colegiaturas, etc.). Dado que los costos deben considerarse en toda decisión de inversión, al igual que los beneficios, el nivel óptimo de la inversión en un hijo en particular es el nivel que maximice el bienestar combinado del hijo, de sus padres y de otros miembros de la familia. Ese nivel variará de una familia a otra dependiendo de factores tales como las aptitudes del hijo, la riqueza de los padres y cuánto amen al hijo sus padres; entre más lo amen, serán menores los costos de la inversión en la crianza, pues, en la medida en que la inversión en la educación del hijo beneficie tanto a los padres como al hijo, porque el bienestar del hijo influye positivamente en la función de utilidad de los padres, el costo de la inversión en la educación del hijo se compensa por una ganancia. Los padres que hacen grandes "sacrificios" por sus hijos no están en peor situación que los padres del mismo ingreso que hacen pocos sacrificios, así como los individuos que gastan una gran parte de su ingreso en vivienda o en viajes no están en peor situación que quienes gastan una fracción menor del mismo ingreso en esas cosas.

Con todo, aun cuando los padres amen mucho a sus hijos, existe un peligro de no invertir lo suficiente. Supongamos que un hijo nace de padres muy

pobres. El hijo tiene un enorme poder potencial de obtener ingresos si se le alimenta, viste, alberga y educa apropiadamente, pero sus padres no pueden darle estas cosas. Esto no importaría si el hijo o los padres pudieran pedir prestado contra la capacidad del hijo para obtener ingresos en el futuro. Pero los costos de la obtención de préstamos contra una corriente de ingresos futuros muy inciertos, y también la imposibilidad de avalar un préstamo con la capacidad de una persona para ganar ingresos (no puede convertirla en su esclavo si no paga), convierten a tales préstamos en un método inviable para el financiamiento de un niño promisorio.

Este problema, aunado al hecho de que algunos padres aman a sus hijos poco o nada, y al mismo tiempo la existencia de un altruismo generalizado hacia los niños en su conjunto (es decir, no sólo hacia nuestros propios hijos), podrían explicar la razón de que se impongan a los padres obligaciones legales de proveer a sus hijos cuidado y sostén, incluida la educación. Otras respuestas sociales, además de las leyes de escolaridad obligatoria, para el problema de la subinversión en el capital humano de los hijos[15] incluyen las leyes del trabajo infantil y la provisión pública y gratuita de educación para los hijos de los pobres. Un argumento a favor de tal subsidio es que las personas educadas son una fuente de beneficios externos. Estas personas son mejores trabajadores, quienes al reducir los costos de las empresas trabajan para generar un excedente del consumidor, y votantes más responsables. Algunas producen ideas nuevas que confieren beneficios que el productor no puede apropiarse por completo. Por ejemplo, las leyes de patentes, al establecer limitaciones de duración y de otros tipos sobre las patentes, pueden evitar que un inventor capte los beneficios sociales plenos de su invento (véase el capítulo XI).

Un grave problema práctico de las leyes que prohíben la negligencia es la decisión de lo que se haría con el niño si la amenaza de multas o encarcelamiento no disuade a los padres de desatenderlo. La respuesta de la ley ha consistido en entregar al niño desatendido a padres sustitutos o a una casa de acogida. Ambas soluciones son poco satisfactorias en vista de la dificultad que existe para monitorear el desempeño del custodio. El Estado no puede *confiar* en los padres sustitutos. En virtud de que no tienen derechos de propiedad sobre los ingresos de por vida del niño, no tienen ningún incentivo económico para hacer la inversión que maximizará tales ingresos.

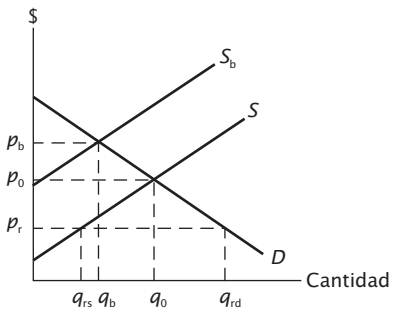
Otra solución para el problema del niño desatendido o indeseado consiste en permitir que los padres (o la madre, si el padre es desconocido o no se interesa) entreguen al hijo para adopción, preferiblemente antes de que em-

[15] Hay una abundante bibliografía económica sobre el capital humano. Véase, por ejemplo, Gary S. Becker, *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education* (3ª ed., 1993). En los capítulos XII, XVII y XVIII volveremos a ocuparnos de este tema. Para los propósitos que convengan en la mayoría de los casos, el capital humano se puede concebir sencillamente como la capacidad de percibir ingresos.

piecen a desatenderlo. La adopción permite que el niño sea transferido de personas que no tienden a invertir óptimamente en su crianza a personas con probabilidades mucho mayores de hacerlo. Pero la anticoncepción eficaz y la declinación del estigma de ser una madre soltera han reducido a casi nada la oferta de niños para adopción. Es posible que los avances recientes en el tratamiento de la fertilidad (quizá promovidos en parte por la disminución de la oferta de bebés para adopción) hayan reducido la demanda de bebés para adopción, pero la demanda sigue siendo mucho mayor que la oferta, en parte porque las mujeres posponen el matrimonio y, por tanto, se encuentran con más problemas de fertilidad. El periodo de espera para conseguir un bebé de una agencia de adopciones se ha alargado hasta varios años, y a veces las agencias no tienen disponibilidad. La escasez de bebés (la cual, sin embargo, se ha mitigado en los últimos años mediante la disponibilidad de las niñas abandonadas en China para su adopción por parte de estadunidenses) se consideraría un ejemplo intolerable de las fallas del mercado si el bien en cuestión fuesen teléfonos celulares en vez de bebés.

En efecto, la escasez parece ser un artificio de la regulación gubernamental, en particular las leyes estatales que prohíben la venta de derechos parentales. Dado que muchas personas que son capaces de tener hijos no quieren criarlos, muchas otras personas que no pueden engendrar a sus propios hijos quieren criar niños, y, dado que los costos de producción para los padres naturales son menores que el valor que muchas personas sin hijos asignan a los niños, sería viable un *mercado* de bebés para adopción. Hay ya un mercado negro de bebés. Su modo de operación necesariamente clandestino impone fuertes costos de información a los participantes en el mercado, así como costos de castigo esperados a los intermediarios (de ordinario abogados y obstetras). El resultado es la existencia de precios mayores y cantidades de ventas menores en relación con lo que ocurriría en un mercado legal.

Esto se muestra en la gráfica V.1, en la que p_0 es el precio del mercado libre de derechos parentales y q_0 es la cantidad del mercado libre, y la regulación gubernamental impone un tope al precio en p_r, muy por debajo de p_0 (p_r no



GRÁFICA V.1.

aparece como cero porque las agencias de adopción y otros proveedores legales cobran honorarios a los padres adoptivos y porque la crianza de los niños es cara). El tope del precio causa una disminución de la cantidad ofrecida a q_{rs}, creando una demanda excedente de $q_{rd} - q_{rs}$. Surge entonces un mercado negro, pero tal mercado es mucho más costoso de operar que un mercado libre (debido a los costos del castigo, la mala información y la carencia de garantías exigibles), y se vacía al precio p_b (donde D se interseca con S_b, la curva de oferta más elevada del mercado negro), que es mayor que el precio del mercado libre. Así, pues, sólo se ofrecen bebés de q_b, en lugar de q_0 al precio del mercado libre.

No todos los bebés se adoptan a través del mercado negro, por supuesto: sólo $qb - q_{rs}$. Las agencias de adopción —organizaciones privadas, no lucrativas, autorizadas por el Estado— utilizan las colas y diversos criterios de fuera del mercado (algunos de ellos muy intrusivos y constitucionalmente cuestionables, como el requisito de que los padres adoptivos tengan la misma religión que los padres naturales) para racionar la inadecuada oferta de bebés que controlan. La objeción principal contra las agencias no es la de los criterios que utilizan para racionar su oferta de bebés, sino su monopolio de las adopciones legales, que asegura (dada su función de ganancia) que la oferta siga siendo inadecuada.

Muchos estados de la Unión Americana permiten también la adopción *independiente* de bebés; los padres naturales (normalmente sólo la madre) arreglan la adopción sin utilizar los servicios de una agencia de adopción. Esto evita los criterios de las agencias, a veces irrelevantes y degradantes, pero, dado que no se permite que la madre *venda* sus derechos parentales, la adopción independiente no crea un mercado real. Sin embargo, se permite que el abogado que arregla la adopción cobre un honorario por sus servicios más el pago de los gastos hospitalarios de la madre y otros relacionados con el nacimiento del niño, y, dado que estos cargos son difíciles de vigilar, en la práctica ocultarán a menudo un pago por el derecho parental mismo. Se aproxima también a una venta simple el "acuerdo familiar". La madre acepta entregar al niño a un familiar cercano a cambio de una retribución financiera para el niño. Tales contratos se han hecho cumplir cuando el tribunal se convenció de que el arreglo beneficiaba al niño.[16]

La respuesta habitual del economista ante un mercado negro consiste en proponer la abolición del control de precios que lo ha generado. Aplicado aquí, esto significaría que se permitiera a las mujeres embarazadas celebrar contratos vinculantes para entregar a su hijo en adopción, sin limitación alguna sobre el precio especificado en el contrato. Se arguye contra esto que no hay ninguna seguridad de que los padres adoptivos que estén dispuestos a pagar más dinero por un niño le proveerán el mejor hogar. Pero las personas que

[16] Enders *vs.* Enders, 30 Atl. 129 (Pa., 1894); Clark *vs.* Clark, 89 Atl. 405 (Md. 1913); In re Estate of Shirk, 350 P.2d 1 (Kan. 1960).

valúan más a un niño tenderán a otorgarle el mayor cuidado,[17] y por lo menos el sacrificio de una suma considerable de dinero para obtener a un niño prueba la seriedad del deseo del comprador de tenerlo. Suponiendo que los padres adoptivos que pagan mucho podrían valuar al niño por malas razones: someterlo al abuso sexual o explotarlo de otro modo, las leyes que prohíben la desatención y el abuso de los niños se aplicarían plenamente a los padres adoptivos (como ocurre por supuesto bajo la ley actual). Podríamos desear que se investigara a los padres adoptivos en cuanto a sus posibles inclinaciones criminales, pero esto ya se hace ahora.

Bien, ¿pero no podría ocurrir que el pago de una gran suma de dinero para obtener los derechos parentales agotara la capacidad financiera de los padres adoptivos para el sostenimiento del niño? ¿O, en términos más realistas, que tal pago redujera su inversión en la crianza del niño? Esto supone, sin embargo, que un mercado libre generaría precios elevados. Ésa es una objeción poco común contra los mercados libres, y está mal aplicada aquí. El precio del mercado no podría superar a los costos de oportunidad (principalmente el tiempo de la madre y los gastos médicos) en los que habrían incurrido los padres adoptivos si hubieran producido al niño ellos mismos en lugar de comprar los derechos parentales sobre él. Porque ése sería el precio competitivo. El costo neto para los padres adoptivos se aproximaría entonces a cero, excepto que los padres adoptivos incurrirían en algunos costos de localización y de determinación de las cualidades del niño en los que no habrían incurrido si hubiesen sido sus padres naturales. El precio del mercado negro es elevado porque debe cubrir los costos del castigo esperado por los vendedores por haber violado la ley y porque la existencia de sanciones legales impide el uso de los métodos más eficientes para la coordinación de vendedores y compradores.[18]

Quienes se oponen al enfoque del mercado sostienen también que los ricos terminarían con todos los bebés, o por lo menos con todos los bebés buenos. (Recuérdese el argumento paralelo contra la venta de frecuencias de radio y televisión.) Eso es improbable que ocurra. Las personas de altos ingresos tienden a tener altos costos de oportunidad del tiempo, y la crianza de

[17] La existencia de un mercado de derechos parentales aumentaría también el incentivo de la madre natural para producir un bebé saludable y reduciría la demanda de abortos.

[18] En cierto sentido, sin embargo, el precio del mercado negro es menor que el que prevalecería en un mercado libre legal. El comprador del mercado negro no recibe ninguna garantía legalmente exigible (de salud, genealogía o lo que sea), comparable a las garantías que reciben los compradores en los mercados legales. El comprador de un mercado legal recibiría un paquete de derechos más valioso, que debiera costar más, porque el vendedor demandaría una compensación por correr los riesgos a que antes se enfrentaba el comprador. Pero el consiguiente aumento del precio sería nominal antes que real; un aumento de precios que sólo compensa la calidad mayor que desea el comprador no reduce el bienestar del comprador. Si giramos a *D* en la gráfica V.1 en sentido contrario a las manecillas del reloj, mostraríamos el efecto de un aumento en la demanda de calidad.

niños es una actividad intensiva en tiempo (incluso los estadunidenses ricos no desean subcontratar la crianza de sus hijos). La demanda total de niños por parte de las parejas ricas sin hijos debe de ser muy pequeña en relación con la oferta de niños que se generaría en un sistema en el que hubiera incentivos económicos para la producción de niños destinados a las parejas sin hijos.

Es posible que a los pobres les vaya peor con la ley de adopción actual que en un mercado libre. Las agencias de adopción, al investigar a los padres adoptivos potenciales a fin de racionar la oferta limitada, otorgan gran importancia al ingreso y el empleo de los solicitantes. Algunas personas que pueden ser rechazadas por los criterios económicos de las agencias podrían adoptar niños en un mercado libre con precios bajos, así como las personas pobres pueden comprar televisores a color. Esta observación puede generalizarse: un grupo desfavorecido (los pobres, los homosexuales, los miembros de una minoría racial, etc.) puede a menudo hacerlo mejor en un mercado privado regulado por el derecho contractual que en las arenas públicas o no lucrativas. Los tribunales que no otorgarían una protección especial a los miembros de tal grupo podrían en cambio hacer cumplir sus contratos.[19] Las empresas lucrativas se preocupan más por el color del dinero de unas personas, y menos por el color de las personas u otras características personales, que las organizaciones no lucrativas.

Aunque la situación en que se encuentra el mercado para la adopción de infantes es una de *demanda* crónicamente excedente, la condición del mercado para la adopción de niños que ya no son infantes es una de *oferta* crónicamente excedente. Un impedimento para su adopción es el hecho de que a los padres sustitutos se les paga por la atención del niño, pero no por su adopción, de modo que si en verdad adoptan a un niño incurrirán en costos mayores aún que los de otros padres adoptivos porque, como sabemos, el ingreso sacrificado es un costo. Este problema podría superarse si se subsidiara la adopción, lo que compensaría la pérdida de ingresos por pasar de padre sustituto a padre adoptivo. El subsidio no tiene que ser tan grande como el pago de cuidado de crianza, ya que el padre adoptivo tiene más derechos sobre el niño: no se le puede quitar tan fácilmente. Esto también significa que la adopción es superior a la acogida temporal debido a que el padre adoptivo, al tener un interés mayor en el niño, invertirá más en su crianza. (Compárese después con XVII.6.)

[19] Como lo destaca Martha M. Ertman, "Contractual Purgatory for Sexual Marginorities: Not Heaven, But Not Hell Either", 73 *Denver U. L. Rev.,* 1107 (1997).

V.5. La maternidad subrogada

La venta de derechos parentales es ilegal en todos los estados de la Unión Americana, pero una práctica estrechamente relacionada —la maternidad subrogada— no lo es. Supongamos que H es fértil, pero M no lo es. H y M contratan a S, una mujer fértil, a fin de que tenga un hijo para la pareja a un precio convenido. S es luego inseminada artificialmente con esperma de H. Cuando nace el niño, S renuncia a sus derechos parentales a favor de M, de acuerdo con el contrato. Ésta es la práctica conocida como maternidad subrogada. En el caso de Baby M,[20] la Corte Suprema de Nueva Jersey sostuvo que el cumplimiento forzoso de los contratos de maternidad subrogada era contrario a la política pública del estado.

La corte expuso varias razones para esta conclusión que demuestran una falta de sofisticación económica y la necesidad consecuente de un libro como éste. "Un niño —dijo la corte—, en lugar de empezar su vida con tanta paz y seguridad como sea posible, se encuentra de inmediato en un campo de batalla entre la madre y el padre contendientes."[21] Pero este campo de batalla es un artificio de la incertidumbre jurídica. Si se arreglara el cumplimiento forzoso de los contratos subrogados, la madre subrogada en el caso de Baby M no habría tenido razones para cuestionar el contrato. "Todo el propósito y efecto del contrato de subrogación era el de otorgar al padre el derecho exclusivo al niño destruyendo los derechos de la madre." Se olvida un hecho obvio: si no hay contrato, no hay niño. El propósito del contrato no era el de extinguir los derechos de la madre sino el de inducir a una mujer a convertirse en madre en beneficio de otra mujer. La corte no entiende la función productiva de los contratos. Cree erróneamente que los contratos sólo reordenan las consecuencias de hechos consumados, como la corte considera el nacimiento de Baby M.

"En la subrogación —continúa la corte—, los mejores postores se convertirán presumiblemente en los padres adoptivos independientemente de la idoneidad." La implicación es la de una oferta fija, tal como ocurre con las pinturas de Van Gogh en una subasta. Pero la oferta no es fija, y es de esperarse que la competencia entre las madres subrogadas potenciales haga bajar el precio hasta el nivel del costo, ubicando la subrogación al alcance de las parejas infértiles de medios modestos. A la corte le preocupan asimismo "los padres adoptivos que pagan más pero no son adecuados". Pero, dado que las personas ricas se ponen invariablemente a la cabeza de la cola en un sistema de adopción regulada, la maternidad subrogada mejorará las perspectivas de las parejas infértiles de medios financieros limitados. "La demanda de hijos es grande y la oferta es pequeña. La disponibilidad de la anticoncepción, el

[20] In re Baby M, 537 A.2d 1227 (N.J., 1988).
[21] Esta cita y las siguientes se han tomado de 537 A.2d en 1247-1249.

aborto y la mayor disposición de las madres solteras para criar a sus hijos ha conducido a una escasez de bebés ofrecidos para adopción. La situación está madura para la entrada del intermediario que traerá cierto equilibrio al mercado aumentando la oferta mediante el uso del dinero." Precisamente. Pero éste es un argumento a favor de los intermediarios, no en contra. La corte culpa a la persona que hace algo acerca de un desbalance entre la demanda y la oferta porque su motivación es financiera. "Es improbable que la subrogación sobreviva sin dinero […] Esa conclusión contrasta con la adopción; por razones obvias, subsiste una oferta sostenida, aunque insuficiente, a pesar de las prohibiciones contra el pago." Ese "aunque insuficiente" es la confesión de que la corte no entiende el sistema de mercado. La oferta es insuficiente *debido* a las prohibiciones contra el pago, y su insuficiencia induce a las parejas infértiles a buscar alternativas en el mercado, tales como la maternidad subrogada. La falla del mercado se encuentra en la adopción, no en la maternidad subrogada.

Michael Sandel, un conocido teórico político, combina la aprobación de la venta de derechos parentales con la condena de los contratos de maternidad subrogada.[22] Un médico de nombre Hicks, que practicaba la medicina en el Sur rural durante los decenios de 1950 y 1960, tenía, según reporta Sandel, "un negocio secreto de venta lateral de bebés". Era también un practicante de abortos que a veces "persuadía a mujeres jóvenes en busca de abortos para que llevaran a sus bebés a término, creando así la oferta que satisfacía la demanda de sus clientes sin hijos". Sandel señala que, comparada con "la empresa casera del doctor Hicks, la subrogación comercial […] es un gran negocio". Pero esto equivale a comparar a un vendedor de un mercado con un mercado entero, y además a un vendedor en un mercado ilegal, donde los vendedores se ocultan, con un mercado legal entero. Dado que hay más de un millón de abortos al año en los Estados Unidos, es evidente el potencial existente de la "venta de bebés", si se legalizara, para eclipsar a la subrogación comercial.

La razón principal de Sandel para distinguir entre la venta de bebés y la maternidad subrogada es que la subrogación, al revés de lo que hacía el doctor Hicks, estimula la comercialización. Señala: "El mercado negro de bebés del doctor Hicks respondía a un problema que surgía independientemente de las consideraciones del mercado. Hicks no alentaba a las madres solteras cuyos bebés vendía para que se embarazaran en primer lugar". No tenía que hacerlo. La demanda incita a la oferta. Las mujeres que sabían de la existencia de un mercado para su bebé si no querían conservarlo tenderían a tener menos cuidado para evitar el embarazo.

[22] Michael J. Sandel, "The Baby Bazaar", *New Republic,* 20 de octubre de 1997, p. 25.

V.6. La ley y la población

Hasta ahora se ha centrado la atención en la calidad antes que en la cantidad de los niños. Pero, por supuesto, el número de niños puede afectar el bienestar de la sociedad en conjunto[23] al afectar el tamaño de la población. En un país densamente poblado, un aumento mayor de la población podría exacerbar problemas ya graves de congestionamiento y de contaminación, imponiendo a otras personas costos que ni el nuevo niño ni sus padres tendrán que pagar jamás. En cambio, en un país escasamente poblado que esté amenazado por enemigos externos, la producción de un niño adicional podría conferir un beneficio por encima de cualesquiera ganancias privadas para el niño y sus padres, al fortalecer (a la larga) al país en lo militar. Y quizá no sean necesarios los enemigos externos para que el crecimiento de la población mejore el bienestar social. Un aumento de la población permite una mayor y más sutil división del trabajo, lo que causará una disminución de los costos y un incremento de la riqueza media y total de la nación.

Por otra parte, un aumento de la población aumenta el incentivo para la innovación tecnológica mediante el aumento del tamaño del mercado para la propiedad intelectual nueva sin el aumento correspondiente en el costo. Por ejemplo, el costo de producir y distribuir un nuevo programa de software no cambia en esencia con el tamaño del mercado, por lo que mientras más grande sea el mercado mayor será el ingreso neto para el productor y, por lo tanto, será mayor el incentivo para crear el software en primer lugar.[24] Sin embargo, al mismo tiempo, el aumento del número de innovadores aumenta la probabilidad de la invención simultánea, lo que reduce las ganancias de la innovación.

Las tasas de natalidad son muy bajas ahora en los países ricos, pero, dado que estos países están también hacinados y que la tecnología militar ha sufrido un viraje en contra de los grandes ejércitos, no está claro que una tasa de natalidad baja —incluso una por debajo del nivel de reposición— debiera causar preocupación. Si fuera así, el problema se rectifica más fácilmente reduciendo las barreras a la inmigración con mayor facilidad que mediante el subsidio de los nacimientos. Un subsidio requerirá impuestos adicionales, lo

[23] El estudio de la población (demografía) es un campo al que los economistas han hecho contribuciones distinguidas desde Malthus, cuya sombría perspectiva llevó a sus contemporáneos a calificar la economía como "la ciencia lúgubre". Véanse tratamientos modernos en, por ejemplo, Gary S. Becker, "An Economic Analysis of Fertility", en su libro titulado *The Economic Approach to Human Behavior,* 171 (1976); Becker y Kevin M. Murphy, "The Family and the State", 31 *J. Law & Econ.*, 1 (1988).

[24] Por ejemplo, si los costos fijos son de $10, el costo variable es de $1 por unidad producida, y el precio es de $2, así, en una producción de 10 unidades, el ingreso total será de $20 y será justamente igual a los costos totales de $20 (10 + [10 × $1]). Pero, en una producción de 20 unidades, el ingreso total será de $40 y superará en $10 los costos totales ($10 + [20 × $1]).

que podría sacar a las mujeres del hogar e introducirlas al mercado en cantidades mayores aún. (Esto depende, con todo, de la forma del impuesto y de otros factores que se discuten en el capítulo XVIII.)

Algunos países asiáticos están tan preocupados por la sobrepoblación que han tratado de limitar cada familia a un número fijo de hijos: en China, a sólo uno (aunque con muchas excepciones y el cumplimiento forzoso laxo de la prohibición restante). Tal enfoque es ineficiente porque las familias difieren en su eficiencia para la producción de hijos. La producción de un segundo hijo podría costarle menos a la familia A de lo que le costaría a la familia B la producción de un primer hijo de la misma calidad. El mismo crecimiento total (positivo o negativo) de la población china podría producirse así a un costo menor otorgando a cada pareja el permiso de un hijo y dando lugar a que los permisos se vendan.

La política demográfica actual de los Estados Unidos es una mezcla confusa, quizás accidental, de subsidios y castigos por la producción de niños. La exención del gravamen del ingreso no pecuniario derivado de la producción hogareña subsidia a la producción de niños en relación con los usos de mercado del tiempo de la esposa, como lo hace por supuesto la exención para los dependientes. La regla común de que los pagos de sostenimiento de los hijos requeridos por una sentencia de divorcio se reducirán si el padre vuelve a casarse y produce hijos del nuevo matrimonio disminuye por debajo del costo social el costo privado, para el padre, de tener más hijos. En la dirección opuesta opera el divorcio fácil, que reduce el incentivo para hacer inversiones maritales a largo plazo, tales como la crianza de hijos, y, por lo tanto, reduce la tasa de natalidad. Pero una ley que prohibiera absolutamente el divorcio podría tener un efecto similar al aumentar la duración óptima de la búsqueda premarital y aumentar así la edad media del matrimonio y reducir el número de los hijos producidos.

Aunque los beneficios y los costos de un aumento en la población varían entre los países, los aumentos poblacionales globales crean externalidades globales cuando el efecto externo, sea positivo o negativo, no se limita geográficamente. Por el lado negativo, mientras haya más seres humanos, y éstos sean más prósperos, se "desbrozará" más la tierra y se quemarán más combustibles fósiles en la combustión interna de los motores y en las plantas de energía eléctrica. Es probable que el desbroce de la tierra reduzca el número de especies, y tanto esto (cuando se hace con fuego, como es común) como la quema de combustibles fósiles provocan emisiones de gases de efecto invernadero y en consecuencia el calentamiento global (véase XIV.9 más adelante). Sin embargo, el aumento de la prosperidad en los países de tercer mundo reduciría la tasa de crecimiento de la población (tal vez a cero), al menos si se correlacionara con un aumento en las oportunidades de mercado para las mujeres, ya que la existencia de tales oportunidades aumenta el costo (de oportunidad) de tener hijos. Además, dado que un porcentaje fijo,

aunque pequeño, de personas que, según se da por sentado, tienen las habilidades y la motivación para diseñar y desarrollar estrategias de destrucción catastrófica, cualquier aumento en la población mundial aumenta el número total de personas nocivas, al igual que mientras haya más planetas hay mayor probabilidad de que en al menos uno de ellos haya vida inteligente. Podrían bastar sólo un puñado de individuos trastornados o fanáticos con habilidades tecnológicas pertinentes para causar una catástrofe global.

¿Cuál es el efecto del aumento de la población en el bienestar humano si las personas adicionales generan una utilidad mayor por estar vivos que la desutilidad que imponen a la población existente? Ésta es una cuestión ética fundamental, ya que la maximización de la utilidad total puede desembocar en una utilidad promedio mínima, mientras que la maximización de la utilidad promedio puede minimizar la utilidad total. Por un lado, imaginemos políticas hipotéticas como las prohibiciones del aborto y la anticoncepción y los enormes subsidios a los hijos que provocaron un aumento significativo de la población, pero una reducción significativa del nivel de vida, y, por otro, políticas que produjeron una disminución sustancial de la población, pero que aumentaron el nivel de vida de manera muy significativa. Sin embargo, si las personas adicionales producidas por las políticas pronatalistas derivan una utilidad positiva de estar vivas sin perjudicar la utilidad de las personas existentes, hay una ganancia neta de utilidad global. No importa si las personas adicionales derivan menos utilidad de estar vivas que la población existente; si en una sociedad de dos personas A tiene una utilidad de 100 por estar viva y B de 90, matar a B no aumentará el bienestar a pesar de que aumente la utilidad promedio de 95 a 100.[25]

El efecto de las políticas pro o antinatalistas en la población puede ser bastante complejo. Por ejemplo, el efecto de la prohibición del aborto en la tasa de natalidad se compensaría parcialmente por el aumento creciente en la demanda de anticonceptivos (sin embargo, no se compensa en su totalidad porque los anticonceptivos no son infalibles y, más importante, a menudo los hombres se niegan a usar condones porque reducen el placer sexual y es probable que las mujeres no tengan acceso a anticonceptivos femeninos eficaces), mientras que las restricciones al aborto que quedaron por debajo de la prohibición total podrían, al elevar el costo del embarazo, reducir la tasa de nacimientos efectivamente.[26] El aborto reduce el costo esperado de las relaciones sexuales (el costo de los hijos no deseados) y, por lo tanto, aumenta la incidencia de las relaciones sexuales. De modo que la fracción de nacimientos no deseados es menor, pero la cantidad de encuentros sexuales, por la que debe multiplicarse esa fracción para obtener el *número* de nacimientos no

[25] David D. Friedman, "What Does 'Optimum Population' Mean?", 3 *Research in Population Econ.*, 273 (1981).

[26] Phillip B. Levine y Douglas Staiger, "Abortion Policy and Fertility Outcomes: The Eastern European Experience", 47 *J. Law & Econ.*, 223 (2004).

deseados, es mayor. Por último, adviértase que, al reducir el número de hijos no deseados, el aborto puede reducir índices futuros de criminalidad, suponiendo que un hijo no deseado recibirá menos inversiones de parte de los padres y, por lo tanto, tendrá perspectivas más pobres de empleo formal cuando crezca.[27]

V.7. ¿Por qué regular el comportamiento sexual?

Sobre todo en las sociedades cristianas y musulmanas, el Estado se encarga no sólo de regular el matrimonio, sino también el comportamiento sexual directamente, o al menos lo intenta aun cuando sea un comportamiento consensual entre adultos. Por ejemplo, muchos estados de la Unión Americana continúan considerando como delitos la fornicación y el adulterio (aunque las leyes se apliquen raras veces); la sodomía era un delito en varios estados hasta que la Corte Suprema declaró inconstitucionales tales leyes; y la prostitución es un delito en todos los Estados Unidos, con excepción de unos cuantos condados de Nevada.

¿Tienen algún sentido económico estas prohibiciones, o son interferencias injustificables con la libertad de contratación?[28] La respuesta podría depender de que las actividades reguladas impongan costos considerables a terceros. Por ejemplo, el adulterio impone comúnmente costos al cónyuge del adúltero. La fornicación podría imponer costos sociales bajo la forma de un niño abandonado o desatendido. Cualquier forma de relación distinta de la relación monógama entre personas que eran vírgenes cuando su relación se inició genera un riesgo de contagio de enfermedades: el sida nos ha vuelto a todos muy conscientes de los peligros de las epidemias de transmisión sexual. Pero, por extraño que pueda parecer, la externalidad generada por las enfermedades de transmisión sexual es menor que en el caso de otras enfermedades contagiosas. La enfermedad de transmisión sexual se difunde primordialmente mediante el contacto *voluntario*, lo que implica (para el economista) que una persona es compensada por asumir el riesgo de contraer la enfermedad. Una pareja sexual potencial y racional exigirá una compensación monetaria o no monetaria por asumir el riesgo de contraer una enfermedad de transmisión sexual. Por lo tanto, el número de casos de enfermedades de transmisión sexual podría aproximarse más al óptimo que en las epidemias habituales transmitidas por el aire, el agua o los insectos. Aproximarse al óptimo, pero no alcanzarlo: la persona que asume voluntariamente un riesgo de infección impone un costo esperado a sus parejas sexuales futuras (y a menudo a su propio futuro y al de sus hijos futuros) sin su consentimiento.

[27] Véase John J. Donohue III y Steven D. Levitt, "The Impact of Legalized Abortion on Crime", 116 *Q. J. Econ.*, 379 (2001).

[28] Véase Richard A. Posner, *Sex and Reason* (1992), esp. caps. 6, 7.

La interrogante más difícil acerca de los efectos de las prácticas sexuales sobre terceros, y una interrogante que el economista no puede contestar, consiste en saber si los costos para el feto abortado, en términos de dolor, pero, aún más importante, en términos de la pérdida de utilidad esperada de la vida, debieran incluirse en el análisis costo-beneficio de los derechos del aborto. No hay nada en la economía que determine quiénes serán incluidos en la sociedad cuyo bienestar habrá de maximizarse. Sin embargo, adviértase que una mujer podría recurrir al aborto porque deseaba posponer la maternidad, y en ese caso el aborto afectará el momento oportuno antes que el número de hijos.

La Corte Suprema ha aprobado restricciones gubernamentales más grandes contra los abortos de embarazos avanzados que contra los abortos realizados en el primer trimestre del embarazo. Mientras más tarde se realice el aborto, más dolor será capaz de sentir el feto y será más semejante a un "niño", un miembro reconocido de la sociedad. No obstante, en contra de este enfoque surge el hecho de que, dado que el aborto es un procedimiento más costoso, causante de mayores tensiones e incluso peligroso para la madre entre más avanzado se encuentre el embarazo al momento de practicarlo, los beneficios del aborto en el embarazo avanzado deberán ser mayores para las mujeres que decidan practicarlo. Adviértase también que, si se prohibiera el aborto en embarazo avanzado, algunas mujeres que temieran tener un feto deforme abortarían más temprano, mientras que si pudieran haber demorado la decisión podrían haber descubierto que el feto era normal y habrían evitado el aborto. (Véase más adelante un argumento paralelo relacionado con el suicidio asistido por un médico en VIII.9.)

Los problemas sociales planteados por el sida, el aborto y los nacimientos fuera de matrimonio (y, por ende, la adopción) se interrelacionan en una forma que el análisis económico puede iluminar.[29] Los condones son eficaces para prevenir la transmisión de la enfermedad; la píldora del control natal no es eficaz en absoluto para este propósito; pero la píldora es más confiable que los condones para la prevención del embarazo. Por lo tanto, era más probable que la epidemia del sida, al aumentar el costo esperado de la enfermedad de transmisión sexual, indujera una sustitución en contra de la píldora del control natal y a favor del condón. Esta sustitución se habría retardado y, por ende, se habría acelerado la difusión de la enfermedad, si la ley hubiera prohibido el aborto. El aborto apoya a la anticoncepción como un método para evitar un nacimiento. Si el apoyo se elimina o se vuelve considerablemente más costoso, aumentará el costo de la anticoncepción imperfecta (por ejemplo, el condón), haciendo más probable que las parejas sexuales escojan un anticonceptivo mejor (por ejemplo, la píldora) antes que el mejor preventivo de la enfermedad.

[29] Véase Tomas J. Philipson y Richard A. Posner, "Sexual Behaviour, Disease, and Fertility Risk", 1 *Risk Decision & Policy*, 91 (1996).

Los conservadores sociales tenían la esperanza de que el sida desalentara el uso de la píldora y de que la restricción del aborto desalentara el del condón, dejando como única alternativa a la abstención: el anticonceptivo perfecto y la mejor prevención de la enfermedad. En términos económicos, los conservadores esperaban que la demanda de sexo se hiciera muy elástica respecto de los riesgos de la enfermedad y el embarazo, de modo que un aumento de tales riesgos provocaría una gran disminución de la cantidad de actividad sexual fuera del matrimonio. La expectativa era poco realista, salvo que la disponibilidad del aborto como apoyo de la anticoncepción sí reduce uno de los costos de las relaciones sexuales (un nacimiento no deseado) y esto debe aumentar la cantidad de relaciones sexuales, lo que debería aumentar la cantidad de enfermedades de transmisión sexual, como los conservadores temen no sin razón.[30] Esto también podría ser resultado de los tratamientos cada vez más eficaces del sida sobre la incidencia de la enfermedad. Mientras más efectivo sea el tratamiento de una enfermedad, será menor el costo total (la muerte y el sufrimiento, así como los costos monetarios) de la enfermedad para el paciente y, por lo tanto, el esfuerzo de la gente para evitar contraer la enfermedad será menor.

Los nacimientos no deseados son un riesgo pequeño de la prostitución y nulo de la relación homosexual, pero la enfermedad de transmisión sexual es un riesgo grande de estas actividades, aunque el peso exacto que a este punto se le debería dar en el diseño de la política económica es, por la razón que ya mencionamos, incierto. En el caso de los bisexuales y también en el de los hombres que frecuentan a las prostitutas, hay un costo externo para los cónyuges… tal vez; porque podría argumentarse que, si estos caminos estuvieran cerrados para los hombres casados, sería más probable que buscaran la variedad sexual en formas más amenazantes para el matrimonio; por ejemplo, las relaciones de largo plazo que podrían conducir al divorcio y a un nuevo matrimonio. Dicho en términos económicos, la prostitución podría ser tanto un sustituto como un complemento de las relaciones maritales, de un modo muy similar a la forma en que la pornografía podría ser un sustituto y un complemento de la violación: lo primero en la medida en que la pornografía vuelve a la masturbación un sustituto más cercano de la relación sexual; lo segundo en la medida en que estimula el deseo de la relación sexual.

Suponiendo que la sociedad desee reducir la cantidad de enfermedades de transmisión sexual, ¿debiera tratar de prohibir la relación promiscua, heterosexual u homosexual? Son inmensos los costos de la represión de delitos sin víctimas (en el sentido de lo consensual) que se cometen en privado. En el caso de las restricciones de las relaciones sexuales motivadas por razones de salud, los costos incluyen el desincentivo que genera la penalización en

[30] Véase Jonathan Klick y Thomas Stratmann, "The Effect of Abortion Legalization on Sexual Behavior: Evidence from Sexually Transmitted Diseases", 32 *J. Legal Stud.*, 407 (2003).

cuanto a la realización de pruebas médicas, la búsqueda de tratamiento, la cooperación en estudios, el aprendizaje de métodos de prevención etc., medidas que podrían ser importantes para el control de una epidemia. Otro costo de la represión, y sin duda mayor, es la disminución de la utilidad causada por disuadir a los individuos de realizar prácticas que disfrutan enormemente; pero esa disminución podría contrarrestarse, por lo menos en parte, por un aumento de la utilidad de los individuos que detestan tales prácticas (véase la siguiente sección de este capítulo).

Las leyes que prohíben la fornicación y el adulterio son quijotescas en la sociedad actual porque los costos de la relación no marital han descendido a un nivel muy bajo. La anticoncepción eficaz ha disminuido los costos del sexo, sobre todo el sexo no marital, que se destina con menos frecuencia a la procreación. Y, en virtud de que las mujeres trabajan cada vez más fuera del hogar, han aumentado para los maridos los costos de su vigilancia, lo que implica una reducción de la probabilidad de la detección. Además, los costos de la búsqueda no marital han bajado porque las mujeres se mezclan con hombres en el trabajo. Con la declinación de los beneficios del matrimonio, más mujeres permanecen sin casarse durante mayor tiempo, de modo que es mayor el fondo de las parejas sexuales no casadas. Incluso los costos de la paternidad singular han bajado para las mujeres, porque ahora las mujeres tienen un ingreso de mercado con el cual pueden comprar los bienes del mercado que necesitan para criar a un hijo. Siendo menos dependientes de los hombres en términos económicos, las mujeres se muestran menos dispuestas a renunciar a su libertad sexual a cambio del sostén económico (los hombres desean este intercambio porque quieren tener derechos sexuales exclusivos sobre la mujer, a fin de garantizar que los hijos que mantienen son propios (de los hombres)). Por lo tanto, disminuye el valor de la virginidad, una demostración de autocontrol y una señal de probable castidad marital.

Cuando las leyes no pueden explicarse como medidas para corregir las externalidades o, de otro modo, para promover la eficiencia económica, el economista se inclina a considerar la probabilidad de que estén diseñadas para redistribuir la riqueza, quizás a instancias de algún grupo de interés. Algunas de nuestras leyes sobre el sexo podrían ser de esta clase. El movimiento antiaborto de la segunda mitad del siglo XIX, que generó las leyes que se derogaron en Roe *vs.* Wade,[31] se vio impulsado en medida considerable por los médicos que, deseando que los profesionales autorizados tuvieran un monopolio de la práctica médica, convirtieron a los practicantes del aborto —que no eran médicos regulares— en el símbolo de los riesgos morales y médicos de permitir que los no regulares proveyeran servicios médicos.[32] Otro ejemplo es la prohibición de la bigamia (poligamia), que al limitar la competencia de

[31] 410 U.S. 113 (1973).

[32] Véase Posner, nota 28, en la página 220.

los hombres por las mujeres aumenta las oportunidades sexuales y maritales de los hombres más jóvenes y más pobres. La prohibición es en efecto un impuesto sobre la riqueza, porque sólo los hombres ricos pueden tener varias esposas. El impuesto no genera una recaudación directamente, pero al disminuir el costo de una esposa transfiere riqueza de los hombres más ricos a los menos ricos (y más numerosos).

V.8. El matrimonio homosexual

En la actualidad, el matrimonio homosexual está autorizado en 15 países y en 13 estados de la Unión Americana además del Distrito de Columbia, aunque las encuestas de opinión pública señalan que más o menos la mitad de la población de los Estados Unidos continúa oponiéndose. La oposición se basa principalmente en la aversión a la idea del matrimonio homosexual. ¿Debiera figurar este rechazo en el cálculo del bienestar social? John Stuart Mill consideraba que no. En *On Liberty* argumentó que un individuo no tiene ningún interés legítimo en las actividades de otras personas que no le afecten, salvo psicológicamente. Mill tenía en mente la indignación que sintieron los ingleses respecto a la poligamia de los mormones, que ocurría a miles de kilómetros de distancia de Utah. Pero ése no es un argumento *económico* porque no existe una diferencia desde un punto de vista de economía del bienestar entre el daño físico y el emocional; cualquiera de ellos reduce la utilidad de la persona perjudicada.

Parte de la oposición al matrimonio homosexual se basa en el temor de que, al volver totalmente respetables las relaciones homosexuales permitiendo tales matrimonios, se fomente la homosexualidad. La mayoría de la gente no quiere que sus hijos sean homosexuales, y esta aversión es un factor en el cálculo de la utilidad. Sin embargo, es casi un hecho que se equivocan al pensar que la homosexualidad se elige; existen pruebas convincentes de que la orientación sexual es una característica innata antes que una característica adquirida. Si es innata no se ve afectada por la aceptación social. Es decir, la *orientación* no se verá afectada; es posible que la *actividad* sí lo sea.

Algunos de los beneficios para las parejas de homosexuales derivados del matrimonio no imponen costos significativos al resto de la población. Por ejemplo, una persona casada no tiene que tener un testamento para legar su propiedad a su cónyuge al morir. Otros beneficios para las parejas casadas imponen costos a terceros; por ejemplo, la seguridad social conyugal y las prestaciones de supervivencia, ya que no las financian en su totalidad los impuestos por seguridad social que pagan las personas que conceden u obtienen los beneficios. Y, recientemente, la Corte Suprema ha argumentado que los cónyuges homosexuales cuyo matrimonio sea válido según la ley del estado

tienen derecho a los mismos beneficios que los cónyuges heterosexuales.[33] Sin duda, el aumento en el costo de los programas de beneficios federales puede tener efectos económicos, pero es poco probable que el aumento derivado del matrimonio homosexual sea grande, ya que al parecer sólo alrededor de 3% de la población es homosexual. Muchos heterosexuales permanecen solteros y sin duda esto también sería cierto para muchos homosexuales, incluso si todos los estados reconocieran el matrimonio homosexual.

Permitir el matrimonio homosexual generaría beneficios sociales compensatorios. Su prohibición aumenta el costo de las relaciones homosexuales monógamas (y, por lo tanto, aumenta la promiscuidad y, por lo tanto, el sida), porque el matrimonio es una contribución a la monogamia. Además, la intolerancia de la homosexualidad, que quizá se reduciría con la aprobación del matrimonio, promueve que los hombres homosexuales se casen con mujeres,[34] lo que aumenta el riesgo de propagación del sida si el marido también tiene relaciones con otros.

Una pregunta un tanto desconcertante es por qué *quieren* el matrimonio los homosexuales, dada la disminución del matrimonio heterosexual y el hecho de que los matrimonios homosexuales son infértiles (aunque, sin duda, eso cambiará con los avances en la ciencia biológica que permiten que el material genético de las personas se mezcle para producir un hijo sin contacto sexual alguno). Los beneficios económicos del matrimonio, como las prestaciones de jubilación de la seguridad social, son un factor, aunque probablemente uno menor. Manifestar el compromiso es un factor mayor. Pero es probable que el factor principal sea el deseo de reconocimiento de las relaciones homosexuales no sólo como legales en el sentido de que no se pueda prohibir la cohabitación de los adultos, sino también como normales (de ahí el término preferido para volver legal el matrimonio homosexual: la "igualdad del matrimonio"). ¿Cómo se sentirían los afroamericanos si un estado les dijera que un hombre negro y una mujer negra pueden cohabitar con todos los derechos de una pareja casada, pero que el nombre que se da a su relación es cohabitación legal (o "unión civil", como algunos estados llaman ahora a lo que otros estados llaman matrimonio homosexual), mas no matrimonio?

Una última y muy difícil pregunta es por qué se ha producido un cambio tan extraordinariamente rápido en las actitudes públicas hacia el matrimonio homosexual. En 1988, 24% de los estadunidenses aprobaban tal matrimonio; en 2013 la cifra había aumentado a 50%, y 67% de las personas lo formaban personas menores de 30 años.

[33] United States *vs.* Windsor, 133 S. Ct. 2675 (2013).

[34] Michael W. Ross, "The Married Homosexual Man: A Psychological Study", 110-111 y cuadro 11.1 (1983). Por ejemplo, el porcentaje de hombres homosexuales que se casan con mujeres en los Estados Unidos es (o más bien era, porque no tenemos estadísticas más recientes que las de Ross) casi el doble que en Suecia. *Idem*.

El crecimiento de los derechos de los homosexuales en general puede atribuirse, en parte, al hecho de que la urbanización creciente ha aumentado el número de homosexuales que, debido a que están geográficamente concentrados, pueden organizarse con mayor eficacia para ejercer la acción política que si se dispersaran. Los homosexuales se concentran en las ciudades en virtud de que el costo del emparejamiento (una forma de los costos de búsqueda) es mayor mientras más escaso sea el mercado en el que se lleve a cabo la búsqueda. Ellos son sólo una pequeña parte de la población, por lo que en una ciudad pequeña o en una zona rural el mercado de parejas homosexuales tiende a ser extremadamente escaso. Pueden reducir el costo del emparejamiento (en concreto, el costo del viaje) si se mudan a las ciudades, donde formarán parte de una población mayor que lo que representan en la nación en conjunto.

Sin embargo, es probable que sean más relevantes para el aumento en la aceptación del matrimonio homosexual los acontecimientos siguientes: una gran aceptación de las relaciones sexuales fuera del matrimonio, y del sexo oral y anal marital o en otras circunstancias, han contribuido a una aceptación creciente del sexo homosexual, que no es marital ni vaginal. El movimiento por los derechos de los gays alentó a los homosexuales a reconocer públicamente su homosexualidad, y a medida que empezaron a "salir del clóset" y se vio que la diferencia mínima respecto a los heterosexuales era sólo la preferencia sexual, los heterosexuales los aceptaron más. Y, por último, luego de que la Corte Suprema considerara que no se puede criminalizar el sexo homosexual entre adultos que dan su consentimiento, y que por lo tanto la cohabitación homosexual era legal, se hizo difícil argumentar en contra de permitir a los cohabitantes la elección del matrimonio.[35]

Lecturas recomendadas

Allen, Douglas W., "An Inquiry into the State's Role in Marriage", 13 *J. Econ. Behavior & Organization,* 171 (1990).

Bailey, Martha J., "'Momma's Got the Pill': How Anthony Comstock and Griswold *v.* Connecticut Shaped US Childbearing", 100 *Am. Econ. Rev*., 98, 104-116 (2010).

Becker, Gary S., *A Treatise on the Family* (edición aumentada, 1991).

———, Elisabeth M. Landes y Robert T. Michael, "An Economic Analysis of Marital Instability", 85 *J. Pol. Econ.,* 1141 (1977).

Brinig, Margaret F., y Douglas W. Allen, "These Boots Are Made for Walking': Why Most Divorce Filers Are Women", 2 *Am. Law & Econ. Rev.,* 126 (2000).

Carbone, June, *From Partners to Parents: The Second Revolution in Family Law* (2000).

[35] Lawrence *vs*. Texas, 539 U.S. 558 (2003).

Economics of Family Law (Margaret F. Brinig [comp.], 2 vols., 2007).
Estin, Ann Laquer, "Love and Obligation: Family Law and the Romance of Economics", 36 *Wm. & Mary L. Rev.,* 989 (1995).
Hadfield, Gillian K., "Households at Work: Beyond Labor Market Policies to Remedy the Gender Gap", 82 *Geo. L. J.,* 89 (1993).
Komura, Mizuki, "Fertility and Endogenous Gender Bargaining Power", 26 *J. Population Econ.*, 943 (2013).
Landes, Elisabeth M., "Economics of Alimony", 7 *J. Legal Stud.,* 35 (1978).
Levmore, Saul, "Love It or Leave It: Property Rules, Liability Rules, and Exclusivity of Remedies in Partnership and Marriage", 58 *Law & Contemp. Prob*., 221 (primavera de 1995).
Lundberg, Shelly, y Robert A. Pollak, "Bargaining and Distribution in Marriage", 10 *J. Econ. Perspectives,* 139 (otoño de 1996).
Matouschek, Niko, e Imran Rasul, "The Economics of the Marriage Contract: Theories and Evidence", 51 *J. Law & Econ.*, 59 (2008).
Mechoulan, Stéphane, "Divorce Laws and the Structure of the American Family", 35 J. Legal Stud., 143 (2006).
Posner, Richard A., *Sex and Reason* (1992).
Rasmusen, Eric, y Jeffrey Evans Stake, "Lifting the Veil of Ignorance: Personalizing the Marriage Contract", 73 *Indiana L. J.,* 453 (1998).
Research Handbook on the Economics of Family Law (Lloyd R. Cohen y Joshua D. Wright [comps.], 2011).
Scott, Elizabeth S., y Robert E. Scott, "Marriage as Relational Contract", 84 *Va. L. Rev.,* 1225 (1998).
Silbaugh, Katharine, "Turning Labor into Love: Housework and the Law", 91 *Nw. U. L. Rev.,* 1 (1996).
Trebilcock, Michael J., y Rosemin Keshvani, "The Role of Private Ordering in Family Law: A Law and Economics Perspective", 41 *U. Toronto L. J.,* 533 (1991).
Tsaoussis, Aspasia, "Protecting Homemakers' Marriage-Specific Investments under No-Fault Divorce: A Model for Restructuring Alimony in Civil Law Countries", 6 *Am. Law & Econ. Rev.*, 217 (2004).

Problemas

1. ¿Puede pensar en una razón económica por la que los esposos tienden a adquirir un seguro de vida mucho mayor que el de las esposas? Si los dos esposos trabajan (en el mercado), ¿debieran tener un seguro de vida total mayor o menor que si sólo trabajara uno de ellos?

2. ¿Podrían los padres invertir en exceso en la educación de sus hijos desde el punto de vista del hijo? ¿Por qué?

3. Leemos en Blackstone: "Por el matrimonio, el marido y la mujer son

una sola persona ante la ley: es decir, el ser mismo o la existencia legal de la mujer se suspende durante el matrimonio, o por lo menos se incorpora y consolida en el del marido". William Blackstone, *Commentaries on the Laws of England,* 442 (1765) (nota de pie omitida). Explique la base económica de esta regla y de su abandono.

4. ¿Hay algún paralelo entre las leyes que prohíben la venta de derechos parentales y las leyes de usura, y entre el mercado negro de bebés y los préstamos usurarios?

5. En un régimen de divorcio incausado, ¿en qué condiciones tenderían más las esposas que los esposos a decidir divorciarse? Véase Margaret F. Brinig y Douglas W. Allen, "'These Boots Are Made for Walking': Why Most Divorce Filers Are Women", 2 *Am. Law & Econ. Rev.,* 126 (2000).

6. Analice el argumento siguiente: Si los padres se divorcian, y se otorga la custodia a uno de ellos, ella (o él) gastará menos en el hijo que lo que gastaría en el matrimonio, porque los gastos de un cónyuge en el hijo confieren un beneficio al otro. Véase Yoram Weiss y Robert J. Willis, "Children as Collective Goods and Divorce Settlements", 3 *J. Labor Econ*., 268 (1985).

7. Se afirma que las mujeres casadas aceptan frecuentemente el sexo con sus maridos sin desearlo realmente. Véase, por ejemplo, Robin L. West, "Legitimating the Illegitimate: A Comment on 'Beyond Rape'", 93 *Colum. L. Rev*., 1442, 1455-1457 (1993). Si la razón es que los hombres tienen en promedio un impulso sexual más fuerte que el de las mujeres, ¿qué sentido *económico* puede dársele a la pregunta retórica del profesor West acerca del matrimonio: "¿Por qué está bien que ella tenga sexo aunque no lo desee, pero no está bien que él no tenga sexo aunque lo desee?"? *Ibid.,* en 1456.

8. Aunque, como se ha señalado en V.2, la tasa de matrimonios en los Estados Unidos se ha reducido drásticamente en los últimos años, la tasa de matrimonios de las mujeres con estudios universitarios, que solía ser sustancialmente menor que la de otras mujeres, en realidad ha aumentado, a pesar de que sigue siendo ligeramente inferior a la de esas otras. ¿Cómo se explica que se haya elevado?

9. Existen pruebas considerables de que las personas casadas son más felices y saludables que las personas solteras. ¿Cómo se explica la diferencia?

VI. EL DERECHO APLICABLE A LOS CUASIDELITOS[1]

VI.1. La economía de los accidentes y la Fórmula de Learned Hand de la responsabilidad por negligencia

Todos toman precauciones contra las lesiones por accidentes; lo interesante es saber cuál es y cuál debería ser la extensión de las precauciones tomadas. Si estuviera considerando la compra de un generador auxiliar para asegurarse de que una falla de la energía no cortará el abasto de oxígeno de su colección de lagartijas sudamericanas estrambóticas, seguramente equilibraría, por lo menos en una forma aproximada y expedita, los beneficios del generador auxiliar en la prevención de la pérdida de las lagartijas contra sus costos. Los beneficios pueden expresarse, por lo menos en una primera aproximación, como el producto de la probabilidad de que las lagartijas mueran dentro de cierto intervalo de tiempo (digamos un año) por una falla de la energía, y la magnitud de la pérdida en dólares (incluida la monetarización de cualquier dimensión emocional). Supongamos que la probabilidad y la magnitud —llamémoslas *P* y *L* (por *loss,* pérdida en inglés)— son 0.001 y 10 000 dólares, respectivamente. Entonces el costo esperado del accidente, *PL*, será de 10 dólares. Es cierto que ésta es una medida de la desutilidad esperada y del costo esperado sólo si usted es neutral al riesgo, pero dejemos a un lado la aversión al riesgo suponiendo que el seguro que tiene sobre las vidas de las lagartijas cubre la muerte por falla de la energía a una prima adicional igual a *PL* más algún cobro moderado (que suponemos igual a cero para simplificar) para cubrir los gastos de administración de la compañía aseguradora. Por lo tanto, el beneficio esperado del generador auxiliar será de 10 dólares anuales para usted. Supongamos que su costo anualizado es de ocho dólares. Entonces comprará el generador siempre que, por supuesto, no disponga de ninguna precaución alternativa más barata. Si el generador cuesta más de 10 dólares, no lo comprará.

Cuando, como en este ejemplo, la persona que toma precauciones y la persona que podría resultar perjudicada si no se toman tales precauciones son la misma, se tomarán precauciones óptimas sin ninguna intervención legal. Pero cambiemos el ejemplo: supongamos que el riesgo no es la pérdida de sus lagartijas sino de su dedo meñique en un accidente automovilístico, y que el procedimiento más barato para evitar el accidente consiste en que algún otro conductor —un total desconocido— conduzca más lentamente. Su-

[1] Véase W. Page Keeton, W. L. Prosser *et al., Prosser and Keeton on Torts* (5ª ed., 1984).

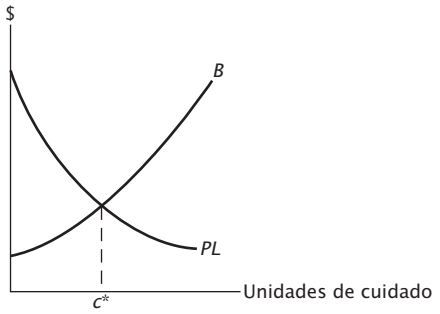
pongamos que, como antes, su costo esperado del accidente es 10 dólares (0.001 × 10 000 dólares), y que el costo de conducir más lentamente para el otro conductor (y así tardarse más en llegar a su destino) es de 8 dólares. La eficiencia requiere que el conductor conduzca más lentamente. Pero, en virtud de que los costos de transacción con las víctimas potenciales tales como usted son prohibitivos, el otro conductor no conducirá más lentamente a menos que el sistema legal intervenga, digamos considerándolo responsable de los daños (= 10 000 dólares) si ocurre un accidente. Por lo tanto, él tiene un costo esperado de 10 dólares por el fallo legal en su contra (si omitimos los gastos legales, el error legal y otras fricciones). Esto lo inducirá a invertir 8 dólares en una precaución que bajará a cero el costo esperado del fallo judicial al prevenir el accidente.

Este ejemplo ilustra la operación del derecho de la negligencia tal como es resumido en la fórmula de la negligencia del juez Learned Hand.[2] Definiendo *P* y *L* como lo hemos hecho aquí, y denotando por *B* el costo de la precaución, Hand escribió que un victimario potencial es negligente si $B < PL$, que es la fórmula implicada por nuestro ejemplo para evitar óptimamente el accidente. Sin embargo, hay una ambigüedad en la formulación de Hand y en la nuestra. Supongamos que aun cuando el *PL* de 10 dólares se eliminaría totalmente si el conductor redujera su velocidad en 40 kilómetros por hora a un costo para él de 8 dólares, podría reducirse a un dólar si el conductor disminuyera su velocidad por sólo 8 kilómetros por hora, a un costo para él de sólo 2 dólares. Esto implica que la disminución de *PL*, de un dólar a cero, le cuesta al conductor 6 dólares (8 – 2), para una pérdida social neta de 5 dólares (6 – 1). Está claro que deseamos que el conductor disminuya su velocidad sólo en 8 kilómetros por hora, lo que genera una ganancia social neta de 7 dólares: la reducción del costo esperado del accidente, 9 dólares, menos el costo de la precaución, 2 dólares.

Este ejemplo revela que los costos esperados del accidente y los costos de la prevención de los accidentes deben compararse en el margen, midiendo los costos y los beneficios de pequeños incrementos de la seguridad y deteniendo la inversión en más seguridad en el punto en el que otro dólar gastado generaría un dólar o menos de seguridad adicional.[3] Afortunadamente, el

[2] Véase United States *vs.* Carroll Towing Co., 159 F.2d, 169, 173 (2d Cir., 1947). Por lo que toca a las aplicaciones contemporáneas de la fórmula, vénase, por ejemplo, Raab *vs*. Utah Ry., 221 P.3d 219, 232 (2009); In re City of New York, 522 F.3d, 279, 285 (2d Cir., 2008); Bourne *vs*. Marty Gilman, Inc., 452 F.3d, 632 (7th Cir., 2006); Mesman *vs*. Crane Pro Services, 409 F.3d, 846 (7th Cir., 2005); Shanklin *vs*. Norfolk Southern Ry. Co., 369 F.3d, 978 (6th Cir., 2004); James *vs*. Meow Media, Inc., 300 F.3d, 683 (6th Cir., 2002); Cross *vs.* Berg Lumber Co., 7 P.3d 922, 936 n.3 (Wyo., 2000); MCI, LLC *vs*. Patriot Engineering & Environmental, Inc., 487 F. Supp. 2d 1029, 1040-1041 (S.D. Ind., 2007).

[3] La Fórmula de Hand marginal se deriva fácilmente con un poco de cálculo diferencial. El problema consiste en encontrar el nivel de cuidado óptimo *(c*),* es decir, el nivel que minimizará el total de los costos esperados de los accidentes, *PL*, más los costos de la prevención de los



GRÁFICA VI.1

método del derecho común facilita un enfoque marginal, simplemente porque de ordinario será difícil para los tribunales la obtención de información sobre cambios no pequeños en las precauciones de seguridad que tome, o que debiera tomar, el victimario.

La Fórmula de Hand, en su forma marginal correcta, se presenta gráficamente en la gráfica VI.1. El eje horizontal representa unidades de cuidado, el eje vertical representa dólares. La curva marcada *PL* describe el cambio marginal de los costos esperados del accidente como una función del cuidado y se muestra declinante bajo el supuesto de que el cuidado tiene un efecto decreciente sobre la prevención de los accidentes. La curva marcada *B* es el costo marginal del cuidado y se muestra creciente bajo el supuesto de que los insumos de cuidado son escasos y, por lo tanto, su precio aumenta a medida que se compran más y más. La intersección de las dos curvas, *c**, representa el cuidado debido. A la izquierda de *c**, el victimario es negligente; *B* es menor que *PL*. A la derecha, donde los costos del cuidado son mayores que los beneficios de la disminución de los costos esperados de los accidentes, el victimario no es negligente; ésta es la región de los accidentes que (sujeto a las calificaciones que discutiremos más adelante) son inevitables en un sentido económico.

Una posible objeción a la Fórmula de Hand es que supone la neutralidad ante el riesgo. Descartamos esta observación señalando que la gente puede comprar seguros, pero en el capítulo IV vimos que las doctrinas del derecho

accidentes, *B*. *P* y *B* son funciones del cuidado del victimario potencial, *c* (estamos suponiendo que las víctimas potenciales no pueden hacer nada para prevenir los accidentes: su cuidado óptimo es cero). Matemáticamente, el problema consiste en minimizar los costos totales, con respecto a *c,* donde *A*(*c*) (los costos totales de los accidentes como una función de la cantidad de cuidado) = *P*(*c*)*L* + *B*(*c*). Siempre que se satisfagan ciertas condiciones bastante naturales, la minimización requiere que se tome la primera derivada de *A* respecto de *c* y se iguale a cero la expresión resultante: *P*c*L* + *B*c = 0, o bien *B*c = – *P*c*L*. En palabras, el cuidado se optimiza cuando un pequeño cambio de *B,* los gastos en cuidado, reduce (de allí el signo negativo) los costos esperados del accidente en la misma cantidad. Éste es el punto *c** en la gráfica VI.1.

contractual no suponen por razones similares que las personas afectadas por ese derecho son neutrales ante el riesgo. Hay una razón económica para esta diferencia. Resulta difícil comprar seguros contra las pérdidas empresariales, porque un empresario asegurado contra la pérdida tendrá escasos incentivos para tratar de prevenirla y porque las pérdidas empresariales son potencialmente como de final abierto; un aumento inesperado del precio de un insumo fundamental podría causar incluso que una empresa grande quebrara. Por lo tanto, los contratos y el derecho contractual desempeñan cierto papel en la disminución del riesgo, lo que ilustra la observación más general de que el seguro de mercado no es la única institución social para la disminución del riesgo (las carteras de inversión diversificadas, que se analizan en el capítulo XVI, son otra). Sin embargo, durante largo tiempo ha habido mercados de seguros bien desarrollados contra las lesiones personales y la muerte, y también contra la responsabilidad por provocar lesiones o la muerte por negligencia. El problema de azar o riesgo moral resultante se atenúa por el hecho de que la mayoría de la gente trata de evitar un accidente, incluso si está completamente asegurada contra las consecuencias para sí misma o para otras personas: la mayoría de las personas son al menos ligeramente altruistas incluso con los extraños y, más importante, a la gente no le gusta lastimarse, incluso si está totalmente asegurada; el aseguramiento puede compensar el dolor pero no eliminarlo. Y ciertamente el dinero no sustituirá por completo la utilidad perdida derivada de una lesión grave, y casi nunca sustituye la utilidad perdida derivada de la muerte (véase VI.12 después). Aun así, el seguro relacionado con los cuasidelitos es muy valorado, y las pérdidas contra las que cubre no son tan impredecibles como para que una compañía de seguros no pueda calcular una prima razonable. De modo que la necesidad de que el sistema de cuasidelitos proporcione un seguro fuera del mercado para remplazar o complementar el seguro de mercado es menos aguda que en el ámbito de los contratos.

Aunque la Fórmula de Hand es relativamente reciente, el método que encapsula se ha utilizado para determinar la negligencia desde que ésta se adoptó por primera vez como el estándar de responsabilidad por regir en los casos de accidentes.[4] Por ejemplo, en Blyth *vs.* Birmingham Water Works[5] se trataba de decidir si una compañía de agua había sido negligente al no enterrar sus tuberías de agua con suficiente profundidad para impedir que reventaran por la congelación y dañaran la casa del demandante. Al fallar a favor de la compañía de agua, el tribunal destacó que la helada había sido de una severidad sin precedente, es decir, que la probabilidad del daño a la propiedad del demandante había sido baja. El daño no era tan grande que hiciese

[4] Véanse Richard A. Posner, "A Theory of Negligence", 1 *J. Legal Stud.*, 29 (1972); Henry T. Terry, "Negligence", 29 *Harv. L. Rev.*, 40 (1915).

[5] 11 Exch. 781, 156 Eng. Rep., 1047 (1856).

al costo esperado del accidente mayor al costo de la prevención, lo que habría implicado un gran gasto para el entierro de las tuberías a mayor profundidad.

Sin embargo, el caso Blyth ilustra un problema en la aplicación de la Fórmula de Hand: el problema de la incertidumbre (véase antes I.1). La Fórmula de Hand supone un riesgo calculable: *P* es un número entre 0 y 1. Pero a menudo no es posible estimar la probabilidad de ocurrencia de un accidente a menos que se tomen algunas precauciones. En el caso de Blyth, la intensidad inusual de la helada sugiere que la probabilidad era baja, pero no podía haber sido posible una mayor precisión. Con todo, dado que *B* era muy alta y *L* sólo era moderada, el hecho de que *P* fuera baja era suficiente para justificar la conclusión de que el acusado no había sido negligente. Si *B* es muy baja y *L* es muy alta, una estimación aproximada de *P* puede ser todo lo que se necesita para justificar la conclusión de la negligencia.

En Adams *vs.* Bullock,[6] un muchacho de 12 años, que cruzaba un puente sobre las vías del trolebús del demandado, agitó un cable de casi dos metros y medio de largo sobre el puente. El cable tocó el alambrado del trolebús del demandado, que corría sobre las vías y por debajo del puente, provocando un choque eléctrico que lesionó al muchacho, quien demandó. El tribunal falló a favor del demandado. *P* era baja porque era improbable que cualquier persona que usara el puente tocara el alambrado del trolebús. Y *B* era elevada; el tribunal, mostrando su perspicacia en materia económica al explicar la diferencia existente entre el alambrado de la luz eléctrica y el de un trolebús, escribió:

> La distinción es que el primero puede ser aislado. La posibilidad del daño, aunque remota [una *P* baja], puede implicar la negligencia, aunque innecesariamente [una *B* muy baja]. La facilidad de la protección puede imponer una obligación de proteger. El caso es diferente con los alambres del trolebús. El aislamiento es imposible. Las guardas acá y allá tienen escaso valor. Para evitar la posibilidad de este accidente y otros parecidos en un punto de la ruta u otro, el demandado debió haber abandonado el sistema instalado y colocar los cables debajo del suelo [una *B* muy elevada].

En Hendricks *vs.* Peabody Coal Co.,[7] un joven de 16 años resultó gravemente lesionado cuando nadaba en la mina abandonada del demandado, que se había llenado del agua de un venero. El demandado, sabedor de que la mina se usaba como un pozo de natación y que era peligrosa a causa de una saliente oculta bajo la superficie en el punto en el que el joven se había tirado un clavado y se había lesionado, instaló vigilancia en el área, pero no

[6] 125 N.E. 93 (N.Y., 1919) (Cardozo, J.).
[7] 253 N.E. 2d 56 (Ill. App., 1969).

lo hizo con eficacia. El tribunal observó que "todo el cuerpo de agua podría haber sido clausurado con una cerca de acero por entre 12 000 y 14 000 dólares. El costo era pequeño en comparación con el riesgo implicado para los niños".

Estos casos demuestran que la aplicación de la Fórmula de Hand depende de la *interacción* de los tres elementos, *B*, *P* y *L*, antes que de cualquiera de ellos tomado por separado. "No se sigue que, cualesquiera que sean las circunstancias, es justificable que se descuide un riesgo de tan pequeña magnitud [baja *P*]. Un hombre razonable sólo descuidaría tal riesgo si tuviera alguna razón válida para hacerlo; por ejemplo, si implicara un gasto considerable la eliminación del riesgo [B alta]. Este hombre ponderaría el riesgo contra la dificultad de su eliminación."[8]

¿Pero qué ocurriría si una precaución beneficiara al victimario potencial al igual que a la víctima potencial? Eso ocurriría en el caso habitual en el que un conductor pone en peligro a otros usuarios del camino: también se estará poniendo en peligro él mismo. ¿Implica esto que la norma del cuidado debiera ser menor, porque ya tiene el conductor un incentivo para evitar un accidente, o mayor, porque el costo esperado del accidente es ahora mayor? Implica lo segundo. En virtud de que los beneficios de evitar un accidente son mayores, el gasto óptimo para evitar el accidente también es mayor.[9] Para ver esto, supongamos que L_v (la pérdida para la víctima si el accidente ocurre) es 100, L_i (la pérdida para el victimario) es también 100, la probabilidad del accidente *(P)* es 0.1, y el accidente puede evitarse a un costo *(B)* de 15. Si ambas *L* se incluyen en la fórmula, el victimario será declarado negligente si no toma la precaución (porque $15 < 0.1\ (200)$), pero si sólo se incluye a L_v no se le declarará negligente (porque $15 > 0.1\ (100)$). En este ejemplo, aunque el victimario potencial se encuentra él mismo en riesgo, éste no es suficientemente grande para inducirlo a tomar la precaución; pero si se le declarara responsable al ocurrir un accidente para otro el victimario tomaría la precaución y el resultado sería un beneficio social neto.

VI.2. El estándar de la persona razonable

Podría pensarse que, si la Fórmula de Hand creara *realmente* los incentivos apropiados para evitar los accidentes por negligencia, nadie sería negligente y no habría juicios por negligencia, por lo menos ninguno ganado por los demandantes, por lo que finalmente no se presentaría ningún caso. Pero por supuesto hay muchos juicios de negligencia ganados por los demandantes, y una explicación es que, al decidir si un accidente pudiera haberse evitado

[8] Wagon Mound (No. 2). Overseas Tankship (U. K.), Ltd. *vs.* Miller Steamship, [1966] All E. R. 709, 718, Privy Council.

[9] Véase Robert Cooter y Ariel Porat, "Does Risk to Oneself Increase the Care Owed to Others?", *Law and Economics in Conflict,* 29 *J. Legal Stud.,* 19 (2000).

por cualquiera de las partes a un costo menor al costo esperado del accidente, los tribunales no tratan de medir las capacidades individuales de las partes para evitar los accidentes. Más bien estiman los costos de evitar los accidentes de la persona típica ("razonable" en la jerga legal). Este enfoque sólo se justifica por los costos de la medición individualizada, que, según se cree, obstaculizan el ajuste necesario. Si la persona típica pudiera evitar a un costo de 120 dólares un accidente cuyo costo esperado era de 100 dólares, pero una persona excepcional pudiera evitar el accidente a un costo menor a 100 dólares, una regla eficiente —sin considerar los costos de la información— impondría una obligación de evitar el accidente a la segunda persona, pero no a la primera. En cambio, la norma de la persona razonable no disuade a la persona excepcional porque no se la hará responsable por provocar el accidente. Si el costo promedio de evitar el accidente fuera de sólo 50 dólares, a la persona que no pueda evitar el accidente a un costo menor de 110 dólares se la haría responsable incluso si no hubiera podido evitar el accidente. Sin embargo, adviértase cómo la regla de la persona razonable disminuye los costos del litigio para el demandante de un cuasidelito. A fin de determinar si tiene una reclamación sólida, no es necesario que este demandante determine la capacidad efectiva del demandado para haber evitado el accidente.

Guardando consistencia con la base del costo de información de la norma de la persona razonable, cuando las diferencias en la capacidad para evitar accidentes pueden determinarse a bajo costo los tribunales reconocen excepciones al estándar de la persona razonable (o a una subclase de tal regla). Por ejemplo, a los ciegos no se les impone un estándar de cuidado tan elevado como a los videntes, aunque dentro de la clase de los ciegos se impone un estándar de cuidado uniforme. Sin embargo, aun cuando no se hace ninguna concesión a las diferencias individuales en la capacidad del cuidado, el estándar de la persona razonable no tiende a producir tasas de error altas. Las personas que incurren en costos de cuidado elevados y en consecuencia son propensas a accidentes todavía pueden protegerse a sí mismas, evitando actividades que las pongan en riesgo; y es poco probable que las personas que incurren en costos de cuidado bajos y, por ende, existen pocas probabilidades de que las hagan responsables de algunos de sus accidentes (accidentes que una persona promedio podría haber evitado) se involucren en accidentes, dada la renuencia normal a involucrarse en accidentes y la mayor capacidad de quien evita accidentes por encima de la media para evitar esa implicación a un costo bajo.

VI.3. La costumbre como una defensa

Una interrogante recurrente es si el cumplimiento del demandado con la norma de seguridad que es habitual en su industria debiera reconocerse como una defensa contra un juicio por negligencia. En caso afirmativo, sólo las

empresas que se encuentren detrás de la empresa típica de su industria, por lo que toca a la adopción de precauciones de seguridad, serán consideradas responsables. Éste es un resultado satisfactorio si hay alguna razón para esperar que la empresa típica tome todas las precauciones justificadas por el costo sin la coerción de la ley. Pero una empresa no tendrá ningún incentivo para tomar precauciones contra los accidentes que son peligrosos sólo para las personas con quienes la empresa no trata ni, por causa de los costos de transacción elevados, puede tratar. Las víctimas potenciales no pagarán a la empresa para que tome precauciones. Tampoco lo harán sus clientes, quienes no se benefician de las precauciones, de modo que, si la empresa tratara de trasladar a ellos los costos añadidos de las precauciones, bajo la forma de precios mayores, se vería desplazada por un competidor. Dado que en estas circunstancias no hay ninguna presunción de que el nivel medio de seguridad de la industria es óptimo, la ley rechaza con razón el cumplimiento de la costumbre como una defensa.

En los casos en los que el accidente es peligroso sólo para los clientes de la industria, es más probable que sea eficiente el nivel habitual de las precauciones tomadas por los vendedores. Los clientes debieran estar dispuestos a pagar precios mayores por el producto o servicio de la industria hasta el punto en que el último dólar gastado compre justo un dólar de disminución del costo del accidente. Por lo tanto, resulta irónico que la enunciación clásica del principio de que el respeto de la costumbre no es una defensa contra una acción de negligencia se haya hecho —¡y por el juez Hand!— en un caso en que el demandante era el cliente del demandado.[10]

En un área de la negligencia, la de la mala práctica médica, los tribunales, congruentes con la distinción que acaba de sugerirse, han permitido tradicionalmente una defensa fundamentada en la costumbre. La obligación de un médico de cuidar a su paciente consiste en cumplir con los estándares habituales de la profesión médica en el área y la especialidad en las que el médico esté ejerciendo su profesión. Los victimarios potenciales (médicos) tienen un incentivo independiente de la ley para proveer el nivel de cuidado por el que las víctimas potenciales están dispuestas a pagar, porque estas últimas son clientes.

Obsérvese el traslape que existe aquí entre los principios de los cuasidelitos y de los contratos. El médico promete implícitamente tratar al paciente con el cuidado habitual entre los médicos del área. Si no emplea todo ese cuidado será responsable de mala práctica, un cuasidelito, pero también ha violado, y por la conducta idéntica, su contrato con el paciente. Sin embargo, en este caso es probable que la responsabilidad de los cuasidelitos sea más eficaz que la responsabilidad de los contratos. Una regla de los cuasidelitos, aplicable a todos los médicos, proporcionará mayor disuasión que los contratos que requieren una redacción cuidadosa, debido a que muchos pa-

[10] The T. J. Hooper, 60 F.2d, 737 (2d Cir., 1932).

cientes no tendrán este tipo de contratos, ya sea porque confían en los médicos o porque los médicos se niegan a asumir voluntariamente la responsabilidad por el cumplimiento negligente.[11]

VI.4. Falta de la víctima: negligencia contribuyente y comparativa, asunción del riesgo y obligaciones para con los allanadores

Que la carga de la precaución sea menor que la probabilidad multiplicada por la magnitud de la pérdida, si no se toma la precaución, es sólo una condición necesaria, no suficiente, para que la precaución sea eficiente. Si otra precaución lograra el resultado deseado a un costo menor aún, ésa sería más eficiente. Dado que, como lo saben todos los peatones, muchos accidentes pueden ser prevenidos por las víctimas a un costo menor que por los victimarios, la ley debe tener cuidado de no perjudicar los incentivos de las víctimas potenciales del accidente para tomar precauciones eficientes. Supongamos que el costo esperado de un accidente de 1 000 dólares podría haberse evitado por el demandado a un costo de 100 dólares, pero también por el demandante a un costo de sólo 50 dólares. La solución eficiente consiste en hacer al demandante "responsable" negándose a permitirle que recupere los daños del demandado. Si el demandado es responsable, el demandante no tendrá ningún incentivo para tomar medidas preventivas porque será plenamente compensado por su perjuicio, y no se obtendrá la solución eficiente.

Hay aquí una exageración, sin duda. Como mencionamos antes, los daños de los cuasidelitos no son siempre plenamente compensatorios, sobre todo cuando están implicados graves perjuicios personales; y, en la medida en que no lo sean, las víctimas potenciales tendrán un incentivo para tomar precauciones, aunque si no las toman no verán disminuido su derecho al pago de daños en un solo centavo en caso de que resulten perjudicados. Pero el incentivo será menor (quizá nulo en los casos de daños a la propiedad, como en el caso de las chispas del ferrocarril que discutimos antes en el capítulo III).

El enfoque tradicional del derecho común, que recibe el nombre de "negligencia contribuyente", era para que el juez o el jurado, después de hallar que el demandado había sido negligente (en caso contrario, ése sería el final del caso), se diera la vuelta y preguntara enseguida si el demandante había sido negligente (y si el accidente se habría evitado en caso de que no lo fuera). Si la respuesta era afirmativa, el demandante perdía. Esto funciona de maravilla en el ejemplo presentado antes, pero supongamos que invertimos las cifras del costo de la precaución, de modo que el costo sea de 50 dólares para el demandado y 100 dólares para el demandante. Parece que el demandado será considerado negligente pero el demandante será contribuyentemente

[11] Jennifer Arlen, "Contracting over Liability: Medical Malpractice and the Cost of Choice", 158 *U. Pa. L. Rev.*, 957 (2010).

CUADRO VI.1

	Súper DC, 0 m	*DC, 22.8 m*	*Sin DC, 61 m*
Cuidado del ferrocarril	$100	$50	$0
Cuidado del agricultor	0	25	110
Costo total	100	75	110

negligente (porque 100 dólares es una cantidad menor que 1 000 dólares) y, por lo tanto, perderá, y que el demandado no tendrá ningún incentivo para tomar en el futuro una precaución que por hipótesis es la precaución eficiente. Esto no es así porque la ley define el cuidado debido como el cuidado que es óptimo *si la otra parte está tomando el cuidado debido*. Dado que, en nuestro ejemplo, si el demandado está tomando el cuidado debido el cuidado óptimo del demandante será cero, el demandante no tendrá ningún incentivo para obrar con cuidado, y sabiendo esto el demandado gastará 50 dólares en cuidado y el accidente se evitará al costo menor. El tribunal no escuchará al demandado si dice que si el demandante estuviera tomando el cuidado debido el cuidado óptimo del demandado sería igual a cero, de modo que él debiera ser exonerado. Porque entonces el victimario potencial no tomaría ningún cuidado, y la víctima potencial evitaría el accidente a un costo más alto: 100 dólares *versus* 50.

Los ejemplos suponen que la solución eficiente era que una de las partes, pero no ambas, obrara con cuidado, de modo que quien tuviera el menor costo para evitar el percance fuera alentado a obrar con cuidado. Pero en ocasiones es más eficiente que cada una de las partes tome cierto cuidado en lugar de que una de las partes obre con cuidado y la otra no haga nada. El cuadro VI.1 ilustra un ejemplo que involucra al cultivo del lino y las chispas de la locomotora. *PL* es 150 dólares. Se determinan los costos de diferentes niveles de cuidado por parte del ferrocarril y del agricultor, y los costos totales se comparan en la tercera hilera del cuadro. *DC* significa guardachispas; un superdetenedor de chispas es simplemente uno de mejor calidad (y más caro). El número de metros es la distancia que media entre los sembradíos de lino y las vías. Entre más alejado se encuentre el lino, más seguro estará; pero hay un costo para el agricultor. Por lo tanto, las tres columnas representan tres combinaciones diferentes de precauciones de seguridad del ferrocarril y del agricultor, todas las cuales tienen los mismos beneficios (150 dólares, el costo esperado del accidente que la combinación evita), pero con costos diferentes.

La combinación de la columna central, donde el ferrocarril instala un guardachispas de calidad media y el agricultor mueve los sembradíos de lino a 22.8 metros de las vías, es la menos costosa, y una regla de negligencia/negligencia-contribuyente inducirá a las partes a adoptarla. Pues supongamos que el ferrocarril no hace nada, esperando que el agricultor sea considerado

negligente porque puede prevenir el accidente a un costo (110 dólares) que es menor que el costo esperado del accidente (150 dólares). Sabiendo que será considerado contribuyentemente negligente sólo por no mantener su lino a la distancia que impedirá su destrucción si el ferrocarril toma la precaución que debe tomar (es decir, si instala el guardachispas), el agricultor ubicará el lino justo a 22.8 metros de distancia de las vías. El lino será destruido, pero, dado que el ferrocarril será considerado negligente y el agricultor no será considerado contribuyentemente negligente, a éste le será indiferente.[12] Sabiendo todo esto, el ferrocarril se verá impulsado a instalar el guardachispas. Es similar el análisis para el agricultor si empieza por ubicar su lino justo al lado de las vías. El ferrocarril no comprará el superguardachispas, porque el agricultor sería considerado contribuyentemente negligente si hubiese un incendio, y la negligencia contribuyente es una defensa completa contra la responsabilidad.

¿Pero *debiera* ser una defensa completa… o, más aún, cualquier defensa? En un régimen de negligencia, si el victimario no es negligente, la víctima soportará todo el costo del accidente, ya sea negligente o no. La defensa de una negligencia contribuyente sólo interviene cuando el victimario es negligente al igual que la víctima. Y si el victimario *es* negligente, ¿por qué debiera ser exonerado, mientras que la víctima soporta todo el costo del accidente? La respuesta económica es que el traslado del costo, de la víctima al victimario, no contribuirá nada a la creación de incentivos para tomar el cuidado debido en el futuro. Ambas partes tienen ya incentivos que en la generalidad de los casos son adecuados. El victimario tiene un incentivo para tomar cuidado a fin de no tener que pagar daños si es descuidado, ocurre un accidente y la víctima no fue descuidada; la víctima tiene un incentivo para tomar cuidado a fin de evitar el accidente si ocurre a pesar de que el victimario fue cuidadoso.

Dado que la eficiencia no aumenta haciendo que el victimario negligente pague daños a la víctima negligente, el derecho común permitía tradicionalmente que el costo del accidente se quedara donde caía, a fin de minimizar los costos de administración del sistema legal. Un pago de transferencia, del victimario a la víctima, costará algo (tan sólo calcular qué tan grande debe ser el pago es una perspectiva abrumadora, como estamos a punto de ver), pero no aumentará la riqueza de la sociedad creando incentivos para el comportamiento eficiente. Sin embargo, la mayoría de las jurisdicciones del derecho común han remplazado la negligencia contribuyente con la negligencia comparativa, por la que si ambas partes (el victimario y la víctima) son negligentes se reduce el pago de daños que debe recibir el demandante, pero no hasta cero.

Sorprendentemente, la negligencia comparativa tiene, al menos en una primera aproximación y sin tomar en cuenta los costos administrativos, los

[12] Éste es un ejemplo de la clase de conducta estudiada en la teoría de juegos. La decisión óptima de cada parte depende de la reacción de la otra parte ante lo que, si no fuese por esa reacción, sería el curso de acción más ventajoso.

mismos efectos que la negligencia contribuyente sobre la seguridad. Regresemos al cuadro VI.1 y supongamos que aun cuando el agricultor no tome ninguna precaución, los daños que el tribunal le otorgará por la pérdida de su cosecha se reducirán sólo en 10%: no en 100%, como en el caso de la negligencia contribuyente. Pareciera que, puesto que el agricultor afronta en consecuencia un costo esperado del accidente de sólo 15 dólares (10% del *PL* de 150 dólares), no gastará los 25 dólares que el cuidado debido requiere que gaste en la precaución. Pero entonces el ferrocarril, sabiendo que no será responsable a menos que sea negligente, pase lo que pase, tendrá todos los incentivos para invertir 50 dólares en precaución a fin de evitar un fallo esperado que le costaría 135 dólares (90% de 150 dólares); y, sabiendo esto, el agricultor tendrá un incentivo para incurrir en el costo de 25 dólares de mover su lino de nuevo a la distancia de 22.8 metros.[13] De otro modo, acabará soportando todo el costo del accidente, porque si el ferrocarril no es negligente no entrará en acción la negligencia comparativa, y la víctima no obtendrá nada. Así se obtiene la asignación de recursos óptima para la seguridad, justo como ocurre cuando la negligencia contribuyente es una defensa completa contra la responsabilidad.

La diferencia económica entre la negligencia contribuyente y la comparativa es que administrar la última es más costoso. Se presentan más demandas y, por lo tanto, más acciones porque la negligencia comparativa imparte un valor esperado positivo a una reclamación que tendría un valor esperado de cero en un sistema de negligencia contribuyente.[14] Y tener que resolver una cuestión adicional (la falta relativa de las partes) introduce incertidumbre adicional en el litigio de un caso de negligencia y, por lo tanto, hace que éste sea más difícil de resolver fuera de los tribunales. La incertidumbre añadida es considerable porque no hay ningún procedimiento objetivo para determinar la falta relativa; éste es un aspecto del problema de la asignación de los costos conjuntos que discutiremos en capítulos posteriores. Aun así, hay una compensación: en un caso cerrado regido por la negligencia contribuyente puede ser bastante incierto si un jurado otorgara una gran sentencia o una sentencia nula, mientras que si el mismo caso se rige por la negligencia comparativa puede ser evidente que el veredicto se encontraría en algún punto intermedio de ese rango. Esta diferencia podría causar que la tasa de acuerdos en los casos regidos por la negligencia comparativa fuera mayor que en los casos regidos por negligencia contribuyente.

Una complicación adicional en la comparación de las dos reglas es que no está claro el efecto de la incertidumbre sobre la cantidad de precaución

[13] No importa el porcentaje por el que disminuyan los daños si se falla que el demandante fue negligente. Los efectos de incentivo son los mismos para cualquier porcentaje entre 0 y 100.

[14] Véase una prueba en Stuart Low y Janet Kiholm Smith, "Decisions to Retain Attorneys and File Lawsuits: An Examination of the Comparative Negligence Rule in Accident Law", 24 *J. Legal Stud.*, 535 (1995).

tomada por victimarios y víctimas, respectivamente.[15] Sin embargo, un estudio empírico a gran escala descubre una conducción menos cuidadosa en las jurisdicciones que aplican la negligencia comparativa.[16] Puede parecer que más demandas y más conducción negligente hacen un argumento convincente en contra de la negligencia comparativa. Sin embargo, es improbable que un desplazamiento de la negligencia contribuyente a la negligencia comparativa induzca menos cuidado por parte de los conductores, ya que la negligencia contribuyente proporciona una defensa absoluta para el demandado negligente y la negligencia comparativa sólo proporciona una defensa parcial.

La negligencia comparativa tiene mayor sentido económico cuando, debido a la dificultad para asegurarse contra los accidentes, la sociedad emplea el sistema de los cuasidelitos para proveer aseguramiento a las víctimas de los accidentes. La negligencia comparativa otorga algo de un victimario descuidado a la víctima descuidada; la negligencia contribuyente no le otorga nada. Por lo tanto, no es sorprendente que la negligencia comparativa obtuviera su primer punto de apoyo en el derecho marítimo. Hasta la época moderna, la transportación marítima era una actividad extraordinariamente riesgosa a causa del gran valor de los barcos y de sus cargas, así como de la probabilidad considerable de desastres, pero el seguro de mercado era difícil de conseguir y en consecuencia había una demanda de aseguramiento por la vía del sistema de los cuasidelitos; ésta se satisfizo mediante varias doctrinas, una de las cuales era la de los daños divididos.[17] La regla en los casos de colisión cuando ambos barcos tenían culpa era que cada parte era responsable de la mitad del daño total sufrido por ambas embarcaciones. El resultado era que el dueño del barco menos dañado asumía una parte de la responsa-

[15] Richard Craswell y John E. Calfee, "Deterrence and Uncertain Legal Standards", 2 *J. Law, Econ. & Organization,* 279 (1986).

[16] Michelle J. White, "An Empirical Test of the Comparative and Contributory Negligence Rules in Accident Law", 20 *RAND J. Econ*., 308 (1989). Véanse también George P. Flanigan *et al.*, "Experience from Early Tort Reforms: Comparative Negligence since 1974", 56 *J. Risk & Ins.,* 525 (1989); Frank A. Sloan, Bridget A. Reilly y Christoph Schenzler, "Effects of Tort Liability and Insurance on Heavy Drinking and Drinking and Driving", 38 *J. Law & Econ*., 49 (1995); Mireia Artigot i Golobardes y Fernando Gómez Pomar, "Contributory and Comparative Negligence in the Law and Economics Literature", en *Tort Law and Economics*, 46, 74-77 (Michael Faure [comp.], 2009). Véase un análisis detallado de las diferencias económicas entre la negligencia contribuyente y la negligencia comparativa en Oren Bar-Gill y Omri Ben-Shahar, "The Uneasy Case for Comparative Negligence", 5 *Am. Law & Econ. Rev.*, 433 (2003).

[17] Otra doctrina interesante, notablemente antigua, lleva el nombre poco informativo de "promedio general". Si se hace necesario que el capitán del barco arroje la carga por la borda, la pérdida se divide entre los fletadores y el dueño del barco de acuerdo con sus intereses relativos en el viaje (es decir, sus participaciones fraccionales en la suma del valor del barco y el valor de las cargas), antes que hacerla recaer sobre el fletador cuya carga se arrojó efectivamente. La regla otorga a cada fletador una forma de seguro sin perjudicar el incentivo del capitán para arrojar la carga que sea más pesada en relación con su valor. William M. Landes y Richard A. Posner, "Salvors, Finders, Good Samaritans, and Other Rescuers: An Economic Study of Law and Altruism", 7 *J. Legal Stud*., 83, 106-108 (1978).

bilidad por el daño causado al otro barco. La regla de los daños divididos tiene la excelente propiedad de proveer un seguro sin alentar al asegurado a ser descuidado, porque si es descuidado y la otra parte de la colisión no lo es pagará el costo total del accidente. Sin embargo, el derecho marítimo se ha desplazado hacia la adopción de un enfoque de la falta comparativa, similar a la negligencia comparativa.

Otra doctrina de la responsabilidad de la víctima es la asunción del riesgo. Se asemeja a la negligencia contribuyente por cuanto es una barrera completa para la recuperación de los daños. Supongamos que una persona entra a una carrera de patines con ruedas plenamente consciente de los riesgos de caídas, y en efecto se cae y se lesiona. Esta persona podría haber sido cuidadosa en el sentido de que, dada la decisión de entrar a la carrera, se condujo como un patinador razonable; y el riesgo al que quedó sujeto podría haber sido sustancial. Supongamos que reduciendo el límite de velocidad en cinco kilómetros por hora podría el dueño de la pista haber evitado todos los accidentes a un costo de tiempo aparentemente trivial para los clientes; sin embargo, el dueño estará protegido contra la responsabilidad ante los clientes lesionados por la defensa de la asunción del riesgo.

Para entender la función económica de la defensa, debemos preguntarnos por qué los clientes no demandan del dueño mayores precauciones de seguridad. Hay varias posibilidades:

1. La Fórmula de Hand se ha aplicado mal. El costo del límite de velocidad menor no es un costo de tiempo trivial sino una disminución sustancial de la emoción de la carrera.
2. La fórmula se ha aplicado demasiado estrictamente. Podría haber sido fácil, para la víctima, evitar el accidente con sólo abstenerse de patinar. El cuidado en la conducción de una actividad es sólo un método de la precaución; otro consiste en evitar la actividad, o en realizarla en menor cantidad (véase la siguiente sección).
3. La velocidad permitida a los patinadores sería un riesgo indebido para la mayoría, pero esta pista atrae a patinadores de agilidad superior. La defensa de la asunción del riesgo permite que personas con capacidades diferentes para evitar el daño se dediquen a actividades de peligrosidad diferente y así introduce un margen de variación en las aplicaciones de la regla de la persona razonable.
4. La pista atrae a quienes prefieren el riesgo. La Fórmula de Hand supone la neutralidad hacia el riesgo y así inducirá a los victimarios potenciales a tomar precauciones que son excesivas para las víctimas potenciales que gustan del riesgo.

Otra regla relacionada de responsabilidad de la víctima es que un terrateniente no es responsable de las lesiones negligentes sufridas por los allana-

dores. En el caso habitual puede prevenirse tal lesión, a un costo menor para el allanador que para el terrateniente, simplemente absteniéndose de allanar. Si el costo de evitarlo es mayor, él puede comprar la tierra (o un derecho de servidumbre o permiso de ingreso) y dejar de ser un allanador. La regla desempeña así la función de alentar las transacciones del mercado, antes que las legales, cuando ello sea viable.

Ocasionalmente no será viable una transacción entre el terrateniente y el allanador, como en el caso de Ploof *vs*. Putnam.[18] El demandante, atrapado en una tormenta, amarró su bote al muelle del demandado. Un empleado de éste alejó al bote, que más tarde fue destruido por la tormenta. El demandante promovió un juicio por la pérdida y ganó. Era grande el valor de allanar la propiedad del demandado durante la tormenta, el costo de prevenir la destrucción era pequeño para el demandado y, en las circunstancias del caso, eran inviables las negociaciones por derechos para atracar el bote.

Pero es probable que el demandante hubiese sido responsable ante el demandado por cualesquiera daños que causara su bote.[19] Tal responsabilidad es apropiada para asegurarse de que el rescate es realmente justificado por el costo, para alentar a los propietarios de muelles a cooperar con los barcos en apuros, para de manera correspondiente desalentar la inversión de éstos en medidas "defensivas" a fin de evitar que un barco en apuros dañe un muelle en un esfuerzo por salvarse a sí mismo[20] (en otras palabras, a fin de obtener la cantidad de inversión en muelles socialmente óptima; véase después VI.9) y, en suma, para simular la inversión y la transacción del mercado que habrían ocurrido si los costos de transacción no hubiesen sido prohibitivos. Adviértase que esto implica que el propietario del muelle también debería tener derecho a un pago por el uso de su muelle. *Cf.* antes IV.15.

En los casos de necesidad pública (Ploof era un caso de necesidad privada), como ocurre cuando el Departamento de Bomberos derriba una casa para combatir un incendio, no se requiere ninguna compensación. Esto ilustra una técnica muy usada por la regulación del derecho común, la de alentar la provisión de beneficios externos (salvando del incendio al resto de la ciudad) permitiendo que los costos (el costo para la persona cuya casa fue derribada) se externalicen.[21] Otro ejemplo es el caso Mouse,[22] citado en el caso Ploof. Mou-

[18] 71 Atl. 188 (Vt., 1908).

[19] Vincent *vs.* Lake Erie Transp. Co., 124 N. W., 221 (Minn., 1910).

[20] Keith N. Hylton, "The Economics of Necessity", 41 *J. Legal Stud.*, 269 (2012).

[21] Durante el gran incendio de Londres de 1666, el alcalde se negó a ordenar que las casas fueran derribadas en la ruta del incendio para formar una brecha, preguntando: "¿Quién pagará el daño?" Richard A. Posner, *Tort Law: Cases and Economic Materials*, 187 (1982). Adviértase que el costo para los propietarios de las casas era bastante menor que el valor de mercado antes del incendio. Las casas estaban a punto de ser consumidas por el fuego, por lo que el costo de oportunidad de su destrucción para crear un cortafuego habría sido pequeño, de hecho cero, si fuera un hecho que el fuego las consumiría.

[22] 12 Co.'63 (King's Bench, 1608).

se era un pasajero en un barco que comenzó a hundirse y cuyos pasajeros tiraron el equipaje por la borda para mantenerlo a flote. Entre los objetos arrojados había un cofre propiedad de Mouse que contenía 113 libras esterlinas. Él demandó al pasajero que había arrojado su cofre y perdió, bajo el argumento de que salvar las vidas de los pasajeros (incluida la de Mouse) valía más que el cofre y su contenido. Es probable que la imposición de una responsabilidad en este caso no había de impedir los rescates (sin duda, el pasajero que tiró el cofre de Mouse por la borda prefirió salvar su vida antes que evitar tener que pagar una sentencia legal), sino que sólo conferiría un beneficio imprevisto para el pasajero cuyo equipaje había vuelto menos seguro el barco.

El demandado en Ploof *vs.* Putnam podría no haber sido considerado negligente si no hubiera habido ningún esfuerzo para sacar el bote del demandante, si el muelle estuviera simplemente en malas condiciones y se derrumbara cuando el demandante trató de amarrar su bote. La probabilidad de que un bote estuviera en peligro en la vecindad del muelle podría haber sido tan pequeña que, de acuerdo con la Fórmula de Hand, el mantenimiento apropiado del muelle no habría sido una precaución justificada por su costo. Pero, en el momento en que el bote del demandante trató de amarrarse, la probabilidad de un accidente grave era elevada, la pérdida esperada del accidente era grande y el costo de evitarlo era pequeño.

Visto así, Ploof *vs.* Putnam es una aplicación especial de la doctrina de la "última oportunidad clara". Un hombre está usando la vía del ferrocarril como un camino. Dado que es un allanador, el ferrocarril no tiene ninguna obligación de cuidarlo (véase antes III.4). Pero si la tripulación lo ve (y advierte que el hombre no se ha percatado de que se aproxima el tren), deberá accionar el silbato del tren y tomar todas las demás precauciones viables para evitar arrollarlo. Aunque el accidente podría haberse evitado a un costo bajo si el allanador hubiese permanecido fuera de la vía, en el momento en que el tren lo está arrollando es el maquinista quien podría evitar un accidente a un costo menor, y este costo es sustancialmente menor que el costo esperado del accidente. Como alternativa, el caso podría verse como uno en el que, aunque el costo de prevención del accidente es para la víctima menor que el costo del accidente, el costo de la prevención del accidente es para el victimario menor aún.

Podría pensarse que, si no existiera la doctrina de la última oportunidad clara, habría menos allanadores —el costo esperado del allanamiento es mayor sin la doctrina de la última oportunidad clara porque el allanador sufre una mayor parte del costo esperado del allanamiento— y, por lo tanto, tal vez, menos accidentes que con la doctrina. Pero esto olvida el carácter probabilístico de la toma de precauciones. Cruzar la línea central de una carretera de dos sentidos es negligente, pero todos lo hacen en alguna ocasión porque sería demasiado costoso adoptar una estrategia de manejo que redujera a cero la probabilidad de cruzarse. Algunas personas que generalmente son cuida-

dosas se encontrarán algunas veces como allanadores. Y, por lo tanto, el hecho de que la doctrina de la última oportunidad clara reduzca un poco el incentivo para no allanar no es una objeción decisiva contra ella, como tampoco el efecto similar de los casos del ganado que invade una propiedad y del perjuicio atrayente —que discutimos en III.5— es una objeción decisiva a aquellas doctrinas que imponen una responsabilidad de los terratenientes hacia los allanadores.

VI.5. La responsabilidad estricta

La responsabilidad estricta de los cuasidelitos significa que alguien que cause un accidente será responsable de los daños de la víctima aunque el perjuicio no pudiera haberse evitado por el ejercicio del cuidado debido (*PL* podría ser 150 y *B* 300 dólares).[23] Como una primera aproximación, la responsabilidad estricta tiene los mismos efectos que la responsabilidad de la negligencia sobre la seguridad, siempre que haya una defensa de negligencia contribuyente, como ocurre de ordinario, aunque a menudo con un nombre diferente. La defensa es, en efecto, *más* importante para la responsabilidad estricta que para la negligencia. En un sistema de negligencia, un victimario potencial siempre puede evitar la responsabilidad no siendo negligente, lo que se supone que cuesta menos que los costos esperados del accidente debido a la negligencia. Según la responsabilidad estricta, el victimario no puede evitar la responsabilidad no siendo negligente, de modo que si no hay una defensa de la negligencia contribuyente incurrirá en costos para evitar los accidentes que la víctima hubiese podido evitar a un costo menor.

Si *B* es menor que *PL*, el demandado estrictamente responsable tomará precauciones para evitar el accidente, así como el demandado en un sistema de negligencia lo hará a fin de reducir sus costos netos. Menos obviamente, si *B* es mayor que *PL*, el demandado estrictamente responsable no tomará precauciones, justo como bajo la negligencia. Es cierto que tendrá que pagar los daños de la víctima. Pero si estos daños, descontados por la probabilidad del accidente, son menores que el costo de evitarlos, el costo esperado de la responsabilidad (= *PL*) será menor que el costo de evitarlos, de modo que esto no es rentable.

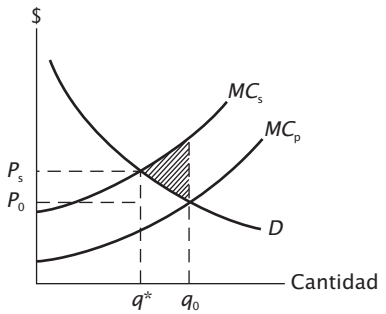
Y sin embargo, hay diferencias económicas considerables entre la negligencia y la responsabilidad estricta. Un procedimiento para evitar un accidente automovilístico consiste en conducir más lentamente; otro procedimiento consiste en conducir menos. En un caso de negligencia, raras veces tratan los tribunales de determinar el nivel óptimo de la actividad que originó el accidente; los tribunales no preguntan si el beneficio del viaje particular (quizás estaba conduciendo hacia la tienda para comprar algo de alimento gour-

[23] Esto es como la responsabilidad por incumplimiento de contrato cuando el incumplimiento es eficiente.

met para su mascota, una iguana) era igual o mayor que los costos, incluido el costo esperado del accidente para otros usuarios del camino, o si manejar era realmente más barato que caminar o tomar el tren cuando se consideran todos los costos sociales. Tal determinación es demasiado difícil para un tribunal en un caso ordinario de cuasidelito. Sólo si los beneficios de la actividad son obviamente muy pequeños —como ocurre cuando un hombre se introduce corriendo a un edificio en llamas para sacar un sombrero viejo y lo hace con el mayor cuidado que puede bajo las circunstancias, pero de todas maneras se quema gravemente— decretará el tribunal que la realización de la actividad era una negligencia en sí misma, aunque, una vez tomada la decisión de realizar la actividad, el actor (demandante o demandado) se haya conducido con la mayor habilidad y circunspección posibles. (Por lo tanto, el caso del sombrero también se podría considerar como un caso de asunción del riesgo.)

Los victimarios potenciales sujetos a una regla de responsabilidad estricta tomarán automáticamente en cuenta los posibles cambios del nivel de la actividad, así como los posibles cambios de los gastos en cuidado, al decidir si prevendrán los accidentes. Supongamos en nuestro ejemplo de las chispas de la locomotora que los guardachispas no funcionan, pero que el ferrocarril podría alcanzar los mismos resultados, a los mismos costos que aparecen en nuestro cuadro VI.1, corriendo menos trenes por día. Éste es un cambio del nivel de la actividad y bien podría pasar inadvertido bajo una regla de negligencia, en cuyo caso el agricultor se vería inducido a tomar todas las precauciones, lo que es una solución inferior.

La gráfica VI.2 muestra el efecto de la responsabilidad estricta sobre la reducción de los costos de los accidentes al inducir cambios en el nivel de la actividad. Se representa la curva de demanda de la industria, porque la regla de la responsabilidad estricta se impone presumiblemente a todos los miembros de la industria. Se supone que los costos de los accidentes inevitables por el cuidado debido y soportados enteramente por extraños a la industria (y, por ende, externos a ella) son proporcionales a la producción de la industria. La curva designada MC_p es la curva del costo marginal privado de la in-



GRÁFICA VI.2

dustria; la curva designada MC_s es la curva del costo marginal social de la industria, es decir, incluye los costos de los accidentes. Según la responsabilidad estricta, MC_s se convierte en la curva del costo marginal privado de la industria, induciéndola a reducir la producción de q_0 a q^*, lo que se traduce en la eliminación de los costos de los accidentes socialmente dispendiosos (el área sombreada en el diagrama).

El problema del empleo de este análisis para apoyar una regla general de responsabilidad estricta es que los cambios del nivel de la actividad realizados por las víctimas son también un método para evitar los accidentes, y un método que se ve alentado por la responsabilidad de la negligencia, pero desalentado por la responsabilidad estricta. Supongamos que el costo para el ferrocarril de la prevención del daño para los cultivos del agricultor, ya sea mediante más cuidado o menos actividad, es mayor a 150 dólares, la pérdida esperada, de modo que el ferrocarril no hará nada, pero que el agricultor podría prevenir la pérdida cambiándose a un cultivo resistente al fuego a un costo de 100 dólares. Según una regla de responsabilidad estricta, el agricultor no tendrá ningún incentivo para hacerlo, porque si no realiza el cambio de su actividad no será considerado negligente contribuyente y, por lo tanto, el ferrocarril tendrá que pagar el daño. Pero según un régimen de responsabilidad por negligencia, dado que el ferrocarril no será responsable por el daño, el agricultor se cambiará al cultivo resistente al fuego, un desplazamiento que ahora le dará una ganancia esperada de 50 dólares.

Sólo si puede identificarse una clase de actividades donde los cambios del nivel de la actividad por parte de victimarios potenciales sean el método más eficiente para la prevención de los accidentes habrá un argumento fuerte para imponer la responsabilidad estricta. En cambio, si hay una clase de actividades en las que los cambios del nivel de la actividad por parte de las víctimas potenciales constituyen el método más eficiente, habrá un argumento fuerte en favor de la no responsabilidad, como en la aplicación de la doctrina de la asunción del riesgo a la participación en deportes peligrosos.

La responsabilidad estricta es la regla para las "actividades ultrapeligrosas": actividades que implican un alto grado de peligro que no puede prevenirse factiblemente por el cuidado del actor o la modificación de su comportamiento por parte de las víctimas potenciales. Tenemos un ejemplo en la responsabilidad estricta por los perjuicios causados por los animales salvajes. Si mi vecino tiene un tigre por mascota, es poco lo que yo puedo hacer (a un costo razonable) para protegerme. Y hay tanto que el dueño puede hacer, siendo cuidadoso, para mantener al tigre bajo control. La precaución más prometedora podría consistir simplemente en que no tenga un tigre: un cambio del nivel de la actividad.[24] Pero supongamos que no estamos hablando

[24] Éste podría parecer un ejemplo ridículo para que lo utilice un residente de Chicago en un libro ostensiblemente serio. Pero en realidad, hace algunos años, los residentes de un grupo de

del tigre de un vecino sino del zoológico. Es improbable que la mejor manera de controlar los accidentes de los visitantes de los zoológicos consista en no tener en esos sitios animales peligrosos: éstos son las atracciones principales del zoológico. Por lo tanto, los tribunales han hecho una excepción a la regla de la responsabilidad estricta en el caso de las lesiones causadas por animales salvajes en los zoológicos, los circos y otros parques y exhibiciones de animales. No es que la responsabilidad estricta persuada efectivamente a los zoológicos para que renuncien a los animales peligrosos. El zoológico compararía el costo de la responsabilidad contra el costo de la pérdida de patrocinio, y el segundo costo dominaría dado que las precauciones que toman los zoológicos para evitar accidentes de los que se les consideraría responsables bajo un estándar de negligencia reducen la probabilidad de un accidente a un nivel muy bajo. Dicho de otra manera, el cuidado es un sustituto superior, en el contexto del zoológico, para la reducción de la actividad (menos animales peligrosos o ninguno).

Otra actividad ultrapeligrosa es la detonación de explosivos. Por cuidadosa que sea la compañía constructora, habrá accidentes; y, dado que la construcción se realiza por todas partes, es improbable que el mejor procedimiento para minimizar estos accidentes sea la modificación de sus actividades por parte de las víctimas potenciales. El mejor procedimiento podría ser que las compañías viraran a métodos alternativos de demolición que sean menos peligrosos; y la responsabilidad estricta crea un incentivo financiero para considerar tales alternativas.

Hay una tendencia a aplicar el rótulo de "actividades ultrapeligrosas" a las actividades nuevas (a menudo llamadas "no naturales"), tales como las presas en Inglaterra[25] o los globos aerostáticos de principios del siglo XIX en los Estados Unidos.[26] Las actividades nuevas tienden a ser peligrosas porque hay escasa experiencia para hacer frente a cualesquiera peligros que presentan. Por la misma razón, los peligros no pueden ser evitables simplemente teniendo cuidado; pero el hecho de que las actividades sean nuevas implica que es probable que haya buenos sustitutos para ellas. Por lo tanto, el mejor método para el control de los accidentes podría ser la reducción de la escala de la actividad (como en la gráfica VI.2), a fin de frenar su difusión mientras se aprende más acerca de cómo emprenderla con seguridad.

La distinción entre la precaución y la actividad no es la única dimensión en la que difieren la responsabilidad por negligencia y la responsabilidad estricta. Otra se relaciona con los costos de administración de estos dos estándares legales. La instrucción de un caso de responsabilidad estricta es más

viviendas situadas a poco más de un kilómetro de la casa del autor pusieron a un león joven en su patio trasero cercado para asustar a los intrusos.

[25] A. W. B. Simpson, "Legal Liability for Bursting Reservoirs: The Historical Context of Rylands *vs.* Fletcher", 13 *J. Legal Stud.,* 209 (1984).

[26] Guille *vs.* Swan, 10 Am. Dec., 234 (N.Y., 1822).

simple que la de un caso de negligencia porque hay un problema menos, el de la negligencia, y mientras menos problemas haya, debiera ser más fácil llegar a un acuerdo sin un juicio; en ambos casos podemos esperar que los costos de litigio sean menores bajo la responsabilidad estricta que bajo la negligencia, para el mismo número de reclamaciones. Pero, como sucede cuando la negligencia comparativa se sustituye por la negligencia contribuyente, es posible que el número no sea el mismo. En principio, bajo la responsabilidad estricta, todo accidente que se traduzca en daños o lesiones puede generar una reclamación, no sólo los accidentes en los que el demandado pudo haber sido negligente.[27] Si la tasa de accidentes en alguna actividad bajara drásticamente al imponerse la responsabilidad estricta, porque los costos de los accidentes superan a los costos de evitarlos mediante cambios en el nivel de la actividad, es posible que haya menos reclamaciones bajo la responsabilidad estricta; y, dado que el costo medio del procesamiento de las reclamaciones debe ser menor bajo la responsabilidad estricta, la sustitución de la negligencia por la responsabilidad estricta será una ganancia económica inequívoca. Pero si la mayoría de los accidentes que ocurren en alguna actividad son inevitables en un sentido económico, ya sea mediante el mayor cuidado o la reducción de la cantidad de la actividad (porque los costos de la mayor precaución o de la menor actividad superan a cualquier ahorro en los costos de los accidentes), el efecto principal del cambio de la negligencia a la responsabilidad estricta será un aumento del número de reclamaciones de daños.

Otra diferencia es que la responsabilidad estricta opera para asegurar a las víctimas de los accidentes inevitables. Pero ésta es una ganancia social sólo si el costo del seguro mediante el sistema de los cuasidelitos es menor que el costo, para las víctimas potenciales, de comprar pólizas de seguros de accidentes en el mercado de seguros; casi seguramente, ese costo es mayor. Todos los participantes en el debate de la no falta convienen en que el sistema de los cuasidelitos es un método muy costoso de provisión del seguro; el debate se centra en la determinación de si ese método provee otro bien, el de la desistencia de los accidentes no justificados por el costo. Véase después VI.14. Una observación relacionada es que, como se muestra en la gráfica VI.2, el tamaño de una industria sujeta a la responsabilidad estricta, y las rentas económicas ganadas en ella, serán menores que si la industria estuviera sujeta a la negligencia.

[27] Además, la responsabilidad estricta se traduce en mayores costos de información para los victimarios potenciales (a diferencia de los costos de la información judicial). En un sistema de negligencia un victimario potencial sólo necesita considerar distintos métodos de cuidado en un sentido estricto. Bajo la responsabilidad estricta también debe considerar la posibilidad de los cambios del nivel de accidente justificados por el costo. En efecto, la responsabilidad estricta desplaza la tarea de la determinación de los métodos óptimos para la prevención de accidentes por parte de los victimarios potenciales de los tribunales hacia los propios victimarios potenciales. Esto sería algo bueno, excepto para los efectos analizados en el texto sobre los incentivos de las víctimas potenciales, y las otras consideraciones exploradas en el texto que hacen que la responsabilidad estricta sea subóptima en una variedad de circunstancias.

Los tribunales cometen errores; ¿cuál régimen —el de la responsabilidad estricta o el de la negligencia— es más fuerte contra los desaciertos? Por una parte, un fallo erróneo en el sentido de que un victimario no es negligente no puede tener consecuencias de mala asignación bajo la responsabilidad estricta, porque la negligencia del victimario no está a discusión. Por otra parte, las consecuencias de una imputación errada de la causación, o una sobrestimación de los daños, son peores bajo la responsabilidad estricta. Bajo la negligencia, una persona es sancionada sólo por su conducta ineficiente; bajo la responsabilidad estricta, puede ser sancionada por una conducta eficiente, y si los costos efectivos de esa conducta se exageran la conducta podría ser disuadida. Supongamos, por ejemplo, que el costo de alguna precaución es 10, el costo esperado del accidente bajo un régimen de responsabilidad estricta libre de error es 9 (y éste es el número que se utiliza para determinar si hay negligencia por no haber tomado la precaución), pero el costo esperado del accidente (en realidad el costo esperado del fallo legal), cuando hay errores acerca de la causación o daños a favor de los demandantes, es 11. Entonces se evitará la conducta bajo la responsabilidad estricta (porque $10 < 11$) pero no bajo la negligencia (porque $10 > 9$), a menos que los errores en cuanto a la causación o los daños infecten la determinación del estándar de la precaución.

A pesar de sus diferencias, sería un error separar a la negligencia y a la responsabilidad estricta. La negligencia tiene un componente de responsabilidad estricta (y en la sección siguiente veremos que la responsabilidad estricta tiene un componente de negligencia). Esto se debe en parte a la regla de la persona razonable, como hemos visto, y en parte a la doctrina de que el superior responde (véase después VI.8). Además, como hemos visto, el cuidado tiene un componente estocástico (es decir, probabilístico). Ser cuidadoso significa tener actitudes, adquirir habilidades y conocimientos, etc., que disminuyan la probabilidad de un error por descuido pero no la eliminen; su eliminación requeriría una inversión excesiva en precaución. Sin embargo, la ley no reconoce la "negligencia óptima", el nivel de cuidado que minimizaría la cantidad de accidentes y los costos de prevención de los accidentes en el largo plazo, pero no necesariamente durante un intervalo corto. Un accidente que ocurre una vez cada diez años a causa de la negligencia puede justificarse por su costo cuando se toma en cuenta el costo de eliminar toda posibilidad de negligencia. Nadie puede ser cauto todo el tiempo sin hacer esfuerzos sobrehumanos que puede no valer la pena intentar.

Se ha sostenido que al ignorar estos puntos la ley crea un sesgo ineficiente a favor de la prevención de accidentes mediante gastos de capital antes que mediante gastos en la operación cuidadosa del equipo e instalaciones existentes.[28] Por ejemplo, supongamos que el costo esperado del accidente es de 10 y que el costo de eliminarlo mediante una mejor capacitación, supervisión, etc.

[28] Mark F. Grady, "Why Are People Negligent? Technology, Nondurable Precautions, and the Medical Malpractice Explosion", 82 *Nw. U. L. Rev.*, 293 (1988).

es de 15. De modo que el accidente ocurrirá, a un costo esperado del juicio legal de 10 para la compañía. Supongamos que una inversión de capital de 5 no eliminaría el accidente, pero eliminaría el error humano en el que se basaría un veredicto de negligencia. Entonces, la empresa tendrá una ganancia neta de 5 (10 – 5) derivada de la inversión, a pesar de que la ganancia social sea de cero.

VI.6. La responsabilidad por los productos

Hoy en día el área más controversial de la responsabilidad estricta es la responsabilidad, ahora llamada estricta en la mayoría de los estados de la Unión Americana, por los perjuicios personales causados (principalmente a los consumidores) por los productos defectuosos o irrazonablemente peligrosos. Pero el término de "responsabilidad estricta" es aquí algo equívoco, porque al decidir si un producto es defectuoso o irrazonablemente peligroso en el diseño o la manufactura los tribunales usan generalmente un enfoque de la Fórmula de Hand, equilibrando los costos esperados del accidente contra los costos de hacer el producto más seguro. Por ejemplo, si un automóvil resulta destruido en un accidente, el fabricante no es declarado estrictamente responsable sólo porque el automóvil no se habría destruido si hubiese sido construido como un tanque. El fabricante sólo es responsable si el diseño o la construcción del automóvil eran defectuosos, lo que en la mayoría de los casos significa casi lo mismo que ser negligente. La excepción principal ocurre cuando, aunque el artículo particular que el consumidor compró era efectivamente defectuoso —una botella de refresco con partes de un ratón en su interior, por ejemplo—, el fabricante había tomado las medidas higiénicas óptimas, y el artículo defectuoso era una falla en un millón de productos que no podría haberse prevenido a un costo menor que el costo (insignificante) esperado del accidente. Sin embargo, el fabricante sería responsable. Ésta es una responsabilidad estricta genuina y puede explicarse por referencia a la discusión de la sección anterior sobre el nivel de la actividad. No hay nada que el consumidor pueda hacer a un costo razonable para impedir una falla en un millón de unidades del producto. No le redituaría inspeccionar cada botella de refresco que compre para prevenir un peligro tan minúsculo, o investigar la posibilidad de que haya sustitutos más seguros; el costo esperado del accidente es demasiado pequeño para incitarlo a tomar cualquier medida de autoprotección. Pero un pequeño costo esperado para cada consumidor, cuando se suma en millones de consumidores, podría ser sustancial. Por hipótesis, imponer este costo al fabricante no lo llevará a tomar medidas higiénicas adicionales en su fábrica; de otro modo sería negligente por no haber actuado ya de ese modo. Pero lo llevará a aumentar su precio;[29] y, aunque el aumento

[29] El costo esperado de las sentencias judiciales, que es un costo marginal porque aumenta

será pequeño (de nuevo por hipótesis), esto hará que algunos consumidores prefieran otros productos que ahora sean más baratos. Si además son más seguros, la actividad consistente en la fabricación y la venta de esta marca de refresco disminuirá ligeramente y con ello el número de los accidentes del producto.

La clase de fallas contra las que los fabricantes garantizan expresamente sus productos son frecuentes y, por ende, familiares para los consumidores, de modo que intervienen en sus decisiones de compra. Pero las fallas de los productos que causan graves perjuicios personales son en extremo raras, y el costo para el consumidor de informarse acerca de ellas tenderá a superar al beneficio esperado. (Si el perjuicio es común y en consecuencia es mejor la información de los consumidores, la defensa de la asunción del riesgo puede impedir que el consumidor perjudicado obtenga el pago de daños, el resultado correcto desde un punto de vista económico.) A menos que el fabricante sea declarado responsable, el mercado podría no disciplinarlo desviando a los consumidores hacia marcas más seguras. La responsabilidad estricta incluye de hecho en su precio la información acerca de los peligros del producto, provocando una sustitución en contra de los productos peligrosos por parte de consumidores que podrían ser completamente ignorantes acerca de los peligros.[30]

Este análisis podría explicar la negativa de la ley a hacer valer las negativas de responsabilidad por los perjuicios personales causados por los accidentes de productos. Si el peligro es muy pequeño, no le convendrá al consumidor prestar atención a la negativa; y, por la misma razón, el fabricante no se verá significativamente afectado cuando aplique la negativa en los casos raros en los que alguien resulta perjudicado. La solución no es perfecta. En particular, si los consumidores difieren en su capacidad para afrontar los peligros de los productos, una negativa de responsabilidad podría ser un procedimiento adecuado para separar a los consumidores de acuerdo con esa capacidad. Los que son más capaces de evitar los riesgos acudirán en masa a los

con la producción, elevará los costos marginales de la empresa y, por tanto, su precio si enfrenta una curva negativa de la demanda o sus competidores también se ven afectados de manera similar por la regla de la responsabilidad. *Cf.* antes III.12.

[30] Este efecto es implícitamente concedido por los oponentes de la responsabilidad estricta en los productos, quienes afirman que ha sacado a algunos productos por completo del mercado. Eso no es necesariamente algo malo porque los costos sociales totales del producto, cuando incluyen los costos del accidente, pueden exceder a los beneficios sociales. Véase Steven P. Croley y Jon D. Hanson, "What Liability Crisis? An Alternative Explanation for Recent Events in Products Liability", 8 *Yale J. Regulation,* 1 (1991). Aun así, hay razones para pensar que la responsabilidad estricta por los productos desalienta la producción de vacunas indebidamente. Éstas tienden a ser productos poco rentables debido a que no se toman de manera repetitiva, como la mayoría de otros fármacos; a menudo nos vacunamos contra una enfermedad determinada una sola vez en la vida. Y, cuando se administra una vacuna a una gran población, cualquiera que desarrolle posteriormente una enfermedad misteriosa tendrá un incentivo para culpar a la vacuna y demandar al fabricante; el número de falsos positivos puede ser grande.

fabricantes que niegan su responsabilidad y cobran los precios menores hechos posibles por sus menores costos de juicios legales. Pero podría haber un procedimiento mejor para separar a los clientes: volviendo obvios los riesgos de un producto mediante una revelación, poniendo así en acción la doctrina de la asunción del riesgo.

El derecho de la responsabilidad por los productos ha ensayado por vez primera una solución interesante para el problema, señalado antes, de comparar los costos de evitar el accidente para el victimario y para la víctima en un caso de cuidado alternativo. Ésta es la doctrina de mal uso previsible y se relaciona con la observación que acabamos de hacer acerca del riesgo obvio. Un fabricante vende una máquina cuyas partes móviles no están protegidas, y un trabajador se lesiona cuando se le atora una mano en ellas. Éste fue descuidado al hacerlo, ya que el peligro era evidente, y sin embargo el fabricante pudo haber protegido las partes móviles, previniendo así el accidente, a un costo trivial. En muchos estados de la Unión Americana, el fabricante habría sido responsable ante el trabajador. Esto es eficiente si el costo es trivial. Recordemos que no es posible reducir a cero la probabilidad de la conducta negligente; por lo tanto, es bueno contar con el apoyo de la responsabilidad si el costo de evitarlo es realmente menor. La cuestión es la misma que en el análisis de la última oportunidad clara, donde la responsabilidad del victimario respalda la incapacidad de la víctima para evitar la posibilidad de que se dañe a sí misma. Pero la doctrina no debiera extenderse al caso del uso negligente ordinario, donde el usuario realmente es quien puede evitar el accidente al menor costo. Eso sería análogo al uso de la doctrina de la asunción del riesgo para eliminar toda responsabilidad bajo la teoría de que cada quien debería saber que existe algún riesgo de ser perjudicado por la negligencia de otro.

VI.7. Causación y previsibilidad

Supongamos que en el caso de Blyth (antes VI.1) la municipalidad no había tomado precauciones contra una helada normal, sino que, como sabemos, la helada que causó daño a la propiedad del demandante era anormal y las precauciones habrían sido inútiles. Dado que el daño causado por la helada anormal no forma parte de *L* en la Fórmula de Hand (siendo *L* la pérdida que habría evitado la precaución que el victimario no tomó), hacer responsable a la municipalidad por el daño causado por la helada sería similar a la imposición de daños punitivos. Si la imposición de daños punitivos por la negligencia simple es un error, como veremos más adelante que lo sería, lo mismo ocurre con la responsabilidad por los daños no causados.

Cambiemos ahora los hechos. A y B, que andan de cacería, descuidadamente toman por error a C por un venado y disparan. Ambos le pegan, y cada disparo es fatal. Esto significa que, vistos por separado, ni A ni B causa-

ron la muerte de C, quien habría muerto de todos modos. Pero sería un error económico dejar libres a ambos, ya que eso reduciría la disuasión. Los herederos de C debieran recibir el pago de daños de A y de B; más adelante discutiremos en qué proporciones. Por extraño que parezca para quien piense que los principios causales deben desempeñar un papel independiente, no relacionado con consideraciones económicas, en la responsabilidad de los cuasidelitos el análisis es el mismo que en el caso que acabamos de ver, aunque sólo una de las balas le pegue a C y no sepamos si A o B la disparó, siempre que tanto A como B hayan actuado descuidadamente.[31]

¿Qué diremos si hay incertidumbre acerca de la identidad de la víctima antes que de la identidad del victimario? Por negligencia, un reactor nuclear emite una oleada de radiación que causará 10 muertes adicionales por cáncer en el barrio del reactor durante los próximos 20 años. Durante este periodo 100 personas habrían muerto de cáncer de todos modos, así que ahora habrá 110; pero nunca sabremos quiénes de los 110 no habrían muerto si no hubiera ocurrido el accidente. Si se considera como causación del cuasidelito un aumento de 10% en la probabilidad de la muerte, entonces cada uno de los 110 podrá recuperar sus daños totales, y el dueño del reactor nuclear se verá obligado a pagar 11 veces el daño que efectivamente causó. Pero si 10% se considera demasiado poco para la causación del cuasidelito, el dueño no pagará nada, y no habrá sanción de cuasidelito por su negligencia. Ninguna de las dos situaciones es satisfactoria. Un enfoque mejor es la definición del perjuicio como el aumento del riesgo de muerte por cáncer, antes que el cáncer mismo. Entonces toda la población expuesta a la oleada de radiación podría demandar cuando el accidente ocurra, empleando el instrumento de la acción de clase (discutida en el capítulo XXII) a fin de economizar en el costo de la litigación de muchas reclamaciones pequeñas. (Adviértase la simetría con el ejemplo anterior, en el que los victimarios potenciales se hicieron responsables como un grupo.)

Ahora consideremos un caso en el que la causación está presente, pero en el que la ley se niega a otorgar el pago de daños. Un tren se descompone a resultas de la negligencia del ferrocarril, y uno de los pasajeros es alojado en un hotel y resulta lesionado cuando el hotel se incendia. Si el tren no se hubiera descompuesto, el pasajero habría continuado hacia su destino y habría dormido en un hotel diferente que no se hubiera incendiado esa noche. Por lo tanto, hay negligencia y una especie de causación, pero no hay responsabilidad. El riesgo de que un hotel se incendie no forma parte de *PL*, el costo esperado del accidente que el ferrocarril podría haber prevenido tomando las precauciones *B*. Es cierto que se habría prevenido el daño de este pasajero particular, pero era igualmente probable *ex ante* que el hotel del pueblo siguiente se hubiera incendiado, en cuyo caso la negligencia del ferrocarril habría conferido un beneficio al pasajero por el que no se habría permitido

[31] El caso principal es el de Sindell *vs.* Abbott Laboratories, 607 P.2d, 924 (Cal., 1980), un caso DES.

que el ferrocarril cobrara. Por lo tanto, declarar responsable al ferrocarril sería imponer (en efecto) daños punitivos por negligencia, justo como en el caso superficialmente muy diferente con el que iniciamos esta sección.

En Rickards *vs.* Sun Oil Co.,[32] la negligencia del demandado puso fuera de servicio al único puente existente entre una isla y la tierra firme. Los comerciantes de la isla, que vieron marchitarse su negocio a resultas de la colisión, demandaron pero perdieron, según la "teoría" de que la pérdida económica pura no es recuperable en un caso de accidente. Aunque los comerciantes de la isla perdieron dinero, su pérdida era una ganancia para los comerciantes de tierra firme que tomaron el negocio de los comerciantes de la isla cuando los clientes ya no pudieron llegar a ella. Dado que el demandado no podría buscar de los comerciantes de tierra firme la restitución de las ganancias que él les había conferido, habría sido punitivo hacerlo pagar las pérdidas de los comerciantes de la isla. El daño causado al puente fue una pérdida social neta, pero recuperable en una demanda promovida por el dueño del puente.

Pero debemos considerar si es probable que la *totalidad* de los beneficios perdidos de los comerciantes de la isla sean capturados por los comerciantes de tierra firme o si algo se escurrirá de las manos y, por lo tanto, habrá una disminución de la riqueza agregada, así como una transferencia de riqueza. No la habrá si no se incrementan los costos para las tiendas de tierra firme derivados de la venta a los clientes adicionales, como no lo serían si, por ejemplo, las tiendas no tuviesen que contratar más personal, o pagar una prima para obtener los bienes adicionales que necesitaran, a fin de satisfacer la creciente demanda; en otras palabras, si las tiendas tuviesen un excedente de capacidad. Aunque el "excedente de capacidad" suena a sobrante, no necesariamente lo es. La mayor parte del tiempo la mayoría de los establecimientos minoristas operan con un poco de excedente de capacidad para poder ser capaces de manejar los picos de la demanda (por lo que un término mejor que "excedente" sería el de la capacidad de "carga máxima"). En el caso Rickards, siempre que el colapso del puente no se produjera en Navidad o en algún otro periodo máximo de venta, las tiendas a las que se desplazaron los clientes de los comerciantes de la isla habrían sido capaces de manejar la demanda adicional a un costo adicional bajo, sobre todo si los clientes se distribuyeran entre un número de almacenes de tierra firme en lugar de amontonarse en uno solo. Sin embargo, si las tiendas de tierra firme carecieran de capacidad no utilizada, el aumento de la demanda generado por el cierre del puente aumentaría los costos de esas tiendas y, por tanto, los precios para todos sus clientes. En este escenario, los comerciantes de la isla y los clientes pierden, y los comerciantes de tierra firme ganan… pero mientras más ganan, más pierden sus clientes al tener que pagar precios altos para compensar a los comerciantes de tierra firme por los costos mayores.

[32] 41 A.2d, 267 (N.J. Sup., 1945).

La dificultad para calcular la pérdida neta en un tipo de caso como el de Rickards hace que sea muy difícil para una empresa que podría precipitar tal pérdida estimar su responsabilidad potencial, por lo que argumenta en contra de la imposición de la responsabilidad. La empresa tendría una idea general de a cuánta gente podría perjudicar, o la cantidad de daños a la propiedad que podría hacer, si el puente colapsara; pero no podría estimar los costos derivados de la interrupción que el colapso pudiera causar en las actividades empresariales y del inconveniente que podría infligir a los clientes de esas empresas. Esta incertidumbre impediría que la empresa determinara cuánto debe invertir en las precauciones para evitar el accidente y podría volver imposible la compra de un seguro de responsabilidad por tales consecuencias. En general, como sabemos, el seguro se puede comprar sólo si la compañía de seguros puede calcular el riesgo de la pérdida, ya que sin dicho cálculo no hay forma de fijar una prima.

A diferencia de la incertidumbre debilitante que enfrenta el victimario potencial, las empresas que podrían verse perjudicadas por tal accidente conocen el valor de su inventario y de sus activos fijos, los comportamientos de sus clientes, el patrón de las demandas, el gasto en personal, etc., y pueden utilizar esa información ya sea para tomar precauciones (el equivalente comercial a abrocharse el cinturón de seguridad) que reduzcan al mínimo las pérdidas de la empresa derivadas de un accidente o para comprar un seguro. Una compañía de seguros debe ser capaz de calcular el riesgo de las pérdidas empresariales por *todos* los accidentes, a diferencia de un accidente provocado por una colisión con un puente determinado.

Es cierto que los seguros de pérdida empresarial están destinados a estar incompletos debido al problema del azar o riesgo moral de asegurar una empresa contra una pérdida de beneficios. Pero la responsabilidad limitada de las empresas que operan en la forma corporativa, y su capacidad para cumplir con sus responsabilidades en caso de quiebra, son, desde un punto de vista funcional, formas de seguro de pérdida empresarial. Y si las víctimas de un cuasidelito económico son numerosas y cada una incurre en sólo una pequeña pérdida, dejar el resto de la pérdida a ellas, antes que desplazarla hacia el victimario, es una forma de seguro automática si bien muy imperfecta. De hecho, es un autoaseguramiento, siendo las víctimas quienes constituyen el fondo de seguros. La mayoría de la gente prefiere autoasegurarse contra el riesgo de pérdidas pequeñas: de ahí los deducibles que se encuentran en la mayoría de las pólizas de seguro, aunque los deducibles tienen la función adicional de inducir al asegurado a ejercer el cuidado, lo que reduce el problema del azar o riesgo moral inherente a los seguros.

El caso Rickards enfatiza la importancia de considerar a la par las ganancias de quien las reciba y las pérdidas de quien las sufra, causadas por un acto presuntamente ilícito. Esto le ayudará a entender, por ejemplo, por qué la competencia no es un cuasidelito. Supongamos que A es el único comer-

ciante de un pueblo pequeño. B acude al sitio y abre una tienda para competir con A. Supongamos que, antes de la aparición de B, el precio que cobraba A por algún bien era de 10 dólares, su ganancia era de 5 dólares y vendía 1 000 unidades, obteniendo así una ganancia total de 5 000 dólares. Después de que B empieza a vender el mismo bien, A se ve obligado a disminuir su precio a 6 dólares. La producción total es ahora, digamos, de 2 000 unidades (mayor que antes porque el precio es menor), y se divide por partes iguales entre A y B, quien vende también a 6 dólares. Las ganancias de A son menores en 4 000 dólares, pero los clientes que habían comprado 1 000 unidades antes y continúan comprándolas mejoran su situación por la misma cantidad, independientemente de que ahora le compren a A o a B. Suponiendo que B sale apenas a mano con el precio de 6 dólares (supongamos que como un recién llegado tiene costos mayores que A), el único otro grupo cuyo bienestar debe ser considerado es el de los consumidores que están comprando el bien por primera vez o comprando más de lo que compraban antes. Todos ellos deben de estar en mejor situación porque de otro modo no se habrían desplazado desde cualquier bien que estuvieran comprando antes. Por lo tanto, hay una ganancia social neta derivada de la nueva competencia, aunque A haya sufrido una pérdida.

Rickards es un emblema de la doctrina de la "pérdida económica" (analizada más adelante con mayor detalle en VIII.2), que distingue entre daños físicos y daños a la propiedad, por un lado, y la "pérdida económica" pura, es decir, una pérdida que no es un daño físico o una pérdida de propiedad, por otro lado, y en general niega la recuperación de las pérdidas económicas sin la compañía de daños físicos o daños materiales. Así que, si una grúa de construcción colapsa delante de una tienda minorista y bloquea la entrada sin dañar la tienda, el propietario no puede demandar por la pérdida de su negocio. (Sin embargo, podría hacerlo si el accidente también hubiese dañado físicamente su tienda. La razón es que, si hay daños físicos, esto limita el número de demandantes potenciales y evita la proliferación de la responsabilidad para las personas que incurran en daños indirectos.) El problema en los casos que hemos visto es que la pérdida económica suele verse compensada por la ganancia económica para otra persona. Pero la doctrina de la pérdida económica también se aplica en una clase de casos en los que esto no es un factor; por ejemplo, cuando un constructor instala ventanas en una casa nueva de manera negligente y el propietario demanda por un cuasidelito debido al costo de volver a instalarlas correctamente.

El fundamento económico de negar una compensación en un caso de este tipo es que los costos de transacción son tan bajos que debiera remitirse al propietario a un remedio de contrato. Pero si se aplica con vigor suficiente la doctrina excluiría la responsabilidad de los cuasidelitos en una serie de casos de daños puramente financieros, tales como el fraude y la mala práctica legal, donde podría parecer que los costos de transacción son tan bajos

como en el caso de la instalación de las ventanas. No se la aplica con tal rigor, porque el fraude y la mala práctica profesional implican una asimetría de información grave entre las partes —el fraude es un encubrimiento y los profesionales tienen una ventaja de conocimiento sobre sus clientes— que hace que un remedio de contrato sea inadecuado para proteger a la víctima potencial.

Un caso que puede parecer incongruente con el de Rickards es White *vs*. Jones:[33] un hombre de 78 años de edad que había sacado de su testamento a sus hijas decidió reintroducirlas e instruyó a sus abogados para que redactaran un nuevo testamento. Los abogados se demoraron sin excusa, de modo que cuando el hombre murió dos meses más tarde el nuevo testamento no había sido redactado todavía. Así que las hijas no obtuvieron nada y demandaron a los abogados… y ganaron. El efecto de la imposición de responsabilidad plena a los abogados por la pérdida causada a las supuestas legatarias fue un aumento de la herencia y, por ende, el otorgamiento de una ganancia súbita a los sobrevivientes del occiso. Porque si el testamento hubiese sido cambiado, como lo deseaba el occiso, los legatarios del testamento anterior habrían obtenido menos, o quizá nada. La diferencia respecto de Rickards es que en este caso la "primera mejor" solución económica —la responsabilidad de los abogados aunada a la restitución de los receptores de la ganancia súbita — habría sido factible. En White, el testamento podría haber sido reformado[34] de tal modo que los legatarios originales habrían tenido que transferir tanto de sus legados como fuera necesario para realizar las intenciones del occiso a los legatarios deseados.

En algunos casos, un demandado escapará a la responsabilidad de las consecuencias de su negligencia porque éstas son imprevisibles. Si esto significara sólo que el accidente había sido improbable y, por ende, inesperado, truncaría arbitraria y drásticamente la responsabilidad del demandado, porque la mayoría de los accidentes son eventos de baja probabilidad. En realidad, significa otras dos cosas. Una es que hay dudas considerables acerca de que la negligencia del demandado fuese una causa del perjuicio. Se le dice a un hombre que su tortuga ha muerto por la negligencia de un veterinario, y al escuchar la noticia cae muerto por un ataque cardiaco; tal consecuencia sería denominada imprevisible, pero lo que parece significar es que no podemos creer que el choque fuese suficiente para matarlo, o pensamos que si se encontraba en un estado tan vulnerable no podría haber vivido mucho tiempo más.

[33] [1995] All E.R. 691 (H.L.).

[34] Por una aplicación de la doctrina de la transferencia fraudulenta, o una analogía de ella. Por ejemplo, si una institución de caridad recibió inocentemente una donación de alguien que había obtenido el dinero de la donación mediante un fraude con valores, la institución tendría que devolver el dinero porque no había hecho consideraciones sobre él, al igual que los beneficiarios originales en White *vs.* Jones no hicieron ninguna consideración sobre sus legados. Véase Scholes *vs.* Lehmann, 56 F.3d, 750 (7th Cir., 1995). Sería un caso de lo que la ley llama "fraude constructivo".

El otro significado de la imprevisibilidad es que los costos elevados de la información impidieron que una de las partes tomara precauciones contra el accidente particular que ocurrió; dicho de otro modo, *B* en la Fórmula de Hand era prohibitiva porque la información acerca del riesgo es un costo de evitar el riesgo. Un allanador se duerme por la noche en una casa recién construida pero desocupada en un desarrollo habitacional y se asfixia porque el constructor había empalmado accidentalmente una manguera de gas y una tubería de agua en la casa. Hay una base económica para negarse a permitir que el constructor interponga la defensa de no obligación de cuidado con los allanadores. A veces, el valor del allanamiento para el allanador es mayor que el costo esperado del accidente (más cualquier daño sufrido por el dueño) y los costos de transacción son prohibitivamente elevados. En estos casos, el allanamiento aumentará el valor. Por lo tanto, deseamos que los allanadores potenciales ponderen los valores y los costos relevantes. Pero tales allanadores no pueden ponderar los costos que son imprevisibles. Un edificio residencial recién construido es normalmente un lugar seguro. El allanador no tiene ninguna razón para prever que se asfixiará. Puede haber razonado de manera perfectamente racional que el valor de su allanamiento superaba todos los costos esperados, incluidos los costos de los accidentes.[35]

Obsérvese la simetría que existe aquí entre los cuasidelitos y los contratos. El allanador asfixiado es como el incumplidor del contrato en Hadley *vs*. Baxendale (véase antes IV.11): ninguno de ellos pudo prever (es decir, informarse al respecto a un costo razonable) las consecuencias de su conducta, y ninguno fue considerado responsable por tales consecuencias. El desarrollador inmobiliario es como el fotógrafo comercial en nuestra variante de Hadley *vs*. Baxendale: cada uno pudo prever las consecuencias de no tomar precauciones y debió ya sea haber tomado precauciones por sí mismo o, si la otra parte podía hacerlo a un costo menor, comunicarle el peligro a la otra parte. Sin embargo, la analogía es imperfecta porque el constructor fue meramente negligente, y no podía haber comunicado un riesgo del que no estaba consciente. Pero hay otra razón para la responsabilidad en el caso del allanamiento: que el allanamiento aumentó la probabilidad del perjuicio. Quien durmiera en la casa resultaría asfixiado, y es circunstancial que fuese un allanador. El allanador podría haber "salvado" al legítimo dueño de morir asfixiado, ¡y, por lo tanto, habría rescatado al constructor de la responsabilidad ante el dueño!

Una aparente anomalía del tratamiento de la previsibilidad en el derecho aplicable a los cuasidelitos es el principio de la "cáscara de huevo" (o del "cráneo delgado"). Se permite a la víctima de un cuasidelito que recupere sus daños plenamente, aunque la *extensión* de su perjuicio era imprevisible para

[35] ¿Debiera el dueño tener derecho a deducir en tal caso un valor de renta justo de cualesquiera daños que se otorgaran al allanador? En principio, sí. Este caso es como el de Vincent *vs*. Lake Erie, que analizamos anteriormente en VI.4.

el demandado a causa de la vulnerabilidad insólita de la víctima. Una razón para imponer a pesar de todo la responsabilidad en tales casos es que, a fin de que el total de los daños del cuasidelito otorgados sea igual al daño total de las víctimas del cuasidelito, debe haber responsabilidad en el caso de la cáscara de huevo para equilibrar la no responsabilidad en el caso del "cráneo de roca" (es decir, cuando la víctima tiene una resistencia al daño por encima del promedio). La alternativa sería otorgar en cada caso los daños en que habría incurrido una víctima media, y este enfoque distorsionaría los incentivos de la víctima potencial: las víctimas potenciales en el caso de la cáscara de huevo tomarían demasiadas precauciones, porque sabrían que no se les compensaría plenamente si resultaran perjudicadas por actividades cuasidelincuenciales, y las víctimas en el caso del cráneo de roca tomarían muy pocas porque serían sobrecompensadas (una ilustración del azar o riesgo moral), y crearían a la vez graves problemas de medición.[36] (Adviértase que los incentivos de los victimarios no se verían afectados a menos que los victimarios pudieran distinguir entre dos tipos de víctima potencial antes de un accidente.)

Estas consideraciones no son igualmente relevantes para Hadley *vs*. Baxendale, porque la existencia de una relación contractual en un tipo de caso como el de Hadley permite que la víctima vulnerable advierta al victimario potencial sobre su vulnerabilidad.[37] Si no lo hace, no puede quejarse de que el victimario potencial no tome en cuenta esa vulnerabilidad. En cambio, la víctima habitual del cuasidelito no tiene ninguna relación contractual con el victimario y, por tanto, no le puede advertir con antelación sobre su vulnerabilidad a fin de inducir mayores precauciones.

VI.8. Cuasidelitos conjuntos, contribución, indemnización; aplicación de la regla de que el superior responde y el hostigamiento sexual

Algunos cuasidelitos los cometen dos o más personas, como cuando dos terratenientes incendian sus propiedades negligentemente y los incendios se extienden, se unen y afectan la casa del demandante, que queda consumida, y cualquiera de los dos incendios podría haber causado todo el daño. Ambos terratenientes, según sabemos, son responsables. O supongamos que un empleado es negligente y perjudica a alguien. El empleador es responsable también, bajo la doctrina de que el superior responde, la que hace a un emplea-

[36] Sin embargo, adviértase una ambigüedad en la idea de los daños plenos en el caso de la cáscara de huevo. Si la cáscara de huevo volvía al demandante vulnerable a otros perjuicios, todos los daños de largo plazo (por ejemplo, por la pérdida de ingresos debido a una incapacidad prolongada) deberán descontarse por la probabilidad de que el demandante hubiera sobrevivido a largo plazo a pesar de su debilidad.

[37] Como lo argumenta David D. Friedman, *Law's Order: What Economics Has to Do with Law and Why It Matters,* 219-220 (2001).

dor responsable de los cuasidelitos de sus empleados, cometidos dentro del campo de su empleo, aunque no haya sido negligente en la contratación, la capacitación, la supervisión o el hecho de no despedirlos. La mayoría de los empleados carecen de recursos para pagar un fallo judicial si perjudican gravemente a alguien y, por lo tanto, no son receptivos a la amenaza de la responsabilidad de los cuasidelitos. El empleador, quien es receptivo, puede inducir a sus empleados a ser cuidadosos, despidiéndolos o castigándolos de otro modo por su descuido. Al hacer responsable a un empleador por los cuasidelitos de sus empleados se le dará un incentivo para que utilice tales inducciones. Adviértase que esta responsabilidad es estricta, lo que es adecuado porque los cambios del nivel de la actividad realizados por el empleador (por ejemplo, sustituyendo insumos de trabajo por insumos de capital o reduciendo la escala de la empresa) son métodos potencialmente eficientes para disminuir los cuasidelitos de los empleados.

La regla de que el superior responde es un ejemplo de castigo colectivo que, aunque a menudo se critica como injusto, es un remedio jurídico convencional, sobre todo en el derecho penal (véase VII.3 más adelante; *cf.* VIII.5 después). La teoría detrás del castigo colectivo es que alguien que no sea el autor de un hecho ilícito podría tener más información que pudiera utilizar (si se le motiva para ello) a fin de evitar el hecho que la que tiene el gobierno. El empleador puede haber sido inocente en el caso particular, pero el saber que de cualquier manera es responsable le dará un gran incentivo para ejercer control sobre sus empleados a fin de prevenir accidentes, incluso mediante medidas indirectas tales como la reducción de su fuerza de trabajo sustituyendo por robots u otros dispositivos mecánicos a los trabajadores humanos falibles.

Sin embargo, el empleador es responsable sólo si el cuasidelito del empleado se cometió en el curso del empleo. Esto significa, por ejemplo, que si el empleado tiene un accidente mientras se dirige o viene del trabajo, el empleador no será responsable. No convendría hacer que el empleador tomara medidas para disminuir los accidentes de sus empleados mientras salen de su trabajo o regresan a él. La probabilidad de tales accidentes es pequeña, y, dado que no ocurren durante el tiempo de trabajo, el empleador carece de información para tomar medidas justificadas por el costo a fin de minimizarlos. Es diferente cuando un empleado que es un vendedor viajero lesiona a alguien mientras hace sus rondas de ventas.

También debe advertirse que el empleador no es responsable por los cuasidelitos de sus contratistas independientes. El contratista independiente tiene mayores probabilidades de ser solvente que un empleado. Además, el empleado recibe un salario por su trabajo bajo la dirección del empleador, mientras que al contratista independiente se le paga un precio por un trabajo contractualmente especificado. El director no supervisa los detalles del trabajo del contratista independiente, de modo que tiene menos probabilidades de estar

capacitado para hacerlo trabajar con seguridad que para hacer que un empleado trabaje con seguridad. Hay, empero, una excepción a la excepción en el caso en que el trabajo del contratista independiente es peligroso. El nivel de la precaución justificada por el costo es por definición mayor, lo que vuelve más probable que el director pueda hacer algo para prevenir los accidentes eficientemente; además, el contratista independiente tiene mayores probabilidades de afrontar un problema de solvencia cuanto más peligroso sea el trabajo y, por ende, cuanto mayor sea el valor esperado de los daños. Sin embargo, aun con esa excepción existe un problema con la regla de que un empleador no es responsable por la negligencia de sus contratistas independientes: le da un incentivo para subcontratar trabajo que los empleados podrían realizar de manera más eficiente.[38]

Ahora se reconoce generalmente al acoso sexual en el lugar de trabajo como un cuasidelito; de modo que podríamos suponer que, si el empleado A acosa al empleado B, su empleador C será responsable ante B. Pero, bajo la visión prevaleciente acerca de este cuasidelito, C será responsable sólo si A fuera un supervisor de B, o si había alguna razón para que C supiera que el acoso sexual era un problema en su lugar de trabajo y no hizo nada al respecto. Así, en el caso del acoso de un cotrabajador visto como distinto del de un supervisor, el empleador es responsable sólo si es negligente; no se aplica aquí la doctrina de que el superior responde. La virtual imposibilidad de que un empleador impida todo acoso sexual de los empleados de bajo nivel privaría a la responsabilidad estricta (por la vía de que el superior responde) de cualquier efecto de asignación benéfico. Pero aun cuando un empleador no pueda controlar todas las interacciones entre sus empleados (a menos que la empresa sea muy pequeña), debe escoger a los supervisores con cuidado y también debiera serle más fácil vigilar su conducta.

Pero, ¿qué es exactamente un "supervisor"? En un caso reciente de acusación de acoso sexual en violación de la ley federal de derechos civiles, la Corte Suprema limitó el término a un empleado facultado para alterar la situación laboral del empleado subordinado que ha sido acosado, por ejemplo, al despedirlo o negarle un bono.[39] La idea es que el empleador ha "proporcionado armas" al "supervisor" para acosar a los empleados que supervisa empoderándolo para tomar represalias en contra de un empleado que rechace sus insinuaciones. Pero esto ignora situaciones en las que el supervisor tiene ese poder sin que el tribunal identifique las armas específicas, como en un caso en el que el supervisor era el único empleado de la compañía en una heladería y sin supervisión estaba a cargo de las chicas menores de edad que eran las únicas otras empleadas.[40] El empleador lo había puesto en una si-

[38] Jennifer H. Arlen y W. Bentley MacLeod, "Beyond Master-Servant: A Critique of Vicarious Liability", en *Exploring Tort Law*, 111 (M. Stuart Madden [comp.], 2005).

[39] Vance *vs*. Ball State University, 133 S. Ct. 2434 (2013).

[40] Doe *vs*. Oberweiss Dairy, 456 F.3d, 704 (7th Cir., 2006).

tuación en la que podía, y en efecto lo hizo, acosar sexualmente a las chicas, a pesar de que no podía despedirlas o reducir sus salarios —aunque, por supuesto, él podría recomendar tales medidas a sus superiores—. En apoyo al enfoque de la Corte Suprema se puede advertir que los mandamases de la empresa suelen supervisar acciones como el despido o la degradación de un empleado, las cuales a menudo requerirán de su aprobación, mientras que si el supervisor utiliza amenazas para forzar una relación sexual con un empleado será extremadamente difícil que los mandamases lo detecten. El caso de la heladería fue intermedio, ya que obviamente alguien por encima del supervisor local había decidido permitir que supervisara a las chicas un solo empleado, que resultó ser un joven.

Tenemos que distinguir entre la responsabilidad de un empleador por los cuasidelitos cometidos por sus empleados y nuestro caso anterior de la responsabilidad de los dos terratenientes por los daños causados por los incendios que habían iniciado. El primer caso se puede calificar como un caso de "cuidado alternativo": el perjuicio se evitará si el empleado se abstiene del acoso o si el empleador lo impide. El segundo es un caso de "cuidado conjunto": para evitar el perjuicio, ambos cuasidelincuentes deben cambiar su comportamiento. La regla tradicional del derecho común en los casos de cuidado conjunto es la de "no contribución entre los cuasidelincuentes", y la regla en los casos de cuidado alternativo es la "indemnización".

La no contribución entre los cuasidelincuentes significa que si el demandante obtiene un fallo contra uno de varios cuasidelincuentes el demandado no puede obligar a los otros a contribuir con ninguna parte del fallo; y si el demandante obtiene un fallo contra un grupo de cuasidelincuentes puede recibirlo de ellos en cualesquiera proporciones que desee, aunque en consecuencia se liberen uno o más de ellos. Podría parecer que esta regla no haría nada para prevenir que se cometa el acto, pero no es así. El análisis es simétrico con el de la negligencia comparativa. Supongamos que hay dos cuasidelincuentes conjuntos potenciales, X y Y. *B* es 10 dólares para X y 20 dólares para Y, y *PL* es 80 dólares. Dado que ésta es una situación de cuidado conjunto, si X o Y no toman las precauciones apropiadas ocurrirá el accidente. Supongamos que X estima que la probabilidad de que se le obligue a pagar los daños es de sólo 10%. Entonces, su costo esperado del accidente es de 8 dólares. Aunque esto es menor que *B*x (10 dólares), X debe saber que si la probabilidad de que él tenga que pagar los daños del demandante es de 10%, la de Y debe ser de 90%, de modo que Y afrontará un costo esperado del accidente de 72 dólares y tendrá un incentivo para gastar 20 dólares en *B*y, en cuyo caso Y no será negligente y todo el costo del accidente recaerá sobre X. Por lo tanto, X tendrá un incentivo para gastar *B*x después de todo. Ambas partes serán cuidadosas.

Aunque la regla de la no contribución es eficiente en los casos de cuidado conjunto, debe ser evidente por la analogía con la negligencia contributi-

vo-comparativa que permitir que un cuasidelincuente conjunto, a quien se obligue a pagar más que su parte "justa" de los daños del demandante, exija la contribución de los otros cuasidelincuentes conjuntos también creará los incentivos de seguridad correctos; y esto independientemente de la forma como se determinen las partes de la contribución (a prorrata, por la falta relativa, etc.). Pero la administración de la contribución es más costosa que la de la no contribución porque requiere que los tribunales decidan otra cuestión y supervisen otro conjunto de pagos de transferencia.

En los casos de cuidado alternativo, el cuasidelincuente conjunto que incurriría en el costo mayor de prevención del accidente puede requerir que el otro cuasidelincuente le rembolse todo el monto de los daños. De este modo, un empleador que se declara responsable bajo la regla de que el superior responde por la negligencia del empleado tiene derecho a una indemnización de parte del empleado por la suma total (aunque esto es rara vez viable) del fallo o por el arreglo que pagó el empleador. De igual modo, el ensamblador final de un producto que ha sido declarado responsable ante un consumidor por los perjuicios causados por un defecto que el ensamblador no podría haber descubierto mediante una inspección razonable tiene derecho a un rembolso de parte del proveedor del componente defectuoso.[41] La explicación económica de estos resultados es que en un caso de cuidado alternativo no queremos que ambos cuasidelincuentes tomen precauciones; queremos que quien pueda evitar el accidente a un costo menor lo haga, ya que eso evitará el accidente. La responsabilidad del otro es un resguardo en caso de que la insolvencia impida que la amenaza de la responsabilidad del cuasidelito disuada al evitador primario del accidente más económico. Así surge la necesidad de un mecanismo que, cuando sea posible, traslade la responsabilidad final a este último, y la indemnización hace esto.

VI.9. Rescate: responsabilidad contra restitución

Mientras camino por la calle, veo que una maceta se cae de una ventana, amenazando a otro peatón, y aunque yo podría salvarlo simplemente gritándole para que se proteja, guardo silencio. El costo esperado del accidente es elevado y el costo de que yo tome la precaución que lo evitaría es trivial, pero no soy responsable. El resultado *parece* inconsistente con la Fórmula de

[41] Éste es otro caso en el que la responsabilidad estricta por los productos es realmente estricta. Este ejemplo, y el de la regla de que el superior responde, sugieren una función económica adicional para la responsabilidad estricta además de las analizadas antes en VI.5: la racionalización del litigio. La víctima demanda al vendedor final o al empleador, según sea el caso, y el demandado puede, mediante la indemnización, trasladar el costo al criminal real si este último es solvente. En este sentido, la responsabilidad estricta se asemeja a los estatutos de "acción directa", que permiten (en algunos estados) que la víctima de un cuasidelito demande directamente a la aseguradora de su victimario.

Hand, ya que si los costos de transacción no hubiesen sido prohibitivos el peatón en peligro me habría pagado seguramente lo suficiente para superar mi renuencia a lanzar un grito de advertencia. Por lo tanto, el hecho de hacerme responsable parecería aumentar el valor. Esta observación se aplica aunque el intento de prevención o de rescate pudiera poner en peligro al rescatador, siempre que el peligro para el rescatador (y, por lo tanto, el costo esperado de las precauciones) sea menor que el peligro para la persona en apuros (y, por lo tanto, el costo esperado del accidente) y que la vida de la víctima sea por lo menos tan valiosa como la del rescatador. Aunque el transeúnte no "causó" el accidente, los conceptos causales desempeñan sólo un papel incidental en el análisis económico de los cuasidelitos. Pero son relevantes en el sentido siguiente. La causación define al conjunto de los demandados potenciales: aquellos que en algún sentido causaron el perjuicio del demandante. Dado que el universo de quienes podrían haber prevenido el perjuicio no está tan circunscrito, habría dificultades prácticas para limitar la responsabilidad del buen samaritano a quienes realmente podrían haber prevenido el perjuicio a un costo razonable.

La imposición de responsabilidad al buen samaritano también volvería más costoso el hecho de encontrarnos en una situación en la que podríamos ser llamados a intentar un rescate, y el costo adicional reduciría presumiblemente el número de los rescatadores potenciales: el nadador fuerte evitaría la playa congestionada.[42] Una objeción a esto es que ello supone que los rescatadores potenciales son también rescatados potenciales. El desincentivo que genera la responsabilidad de estar en un lugar donde podríamos ser responsables por no rescatar a una persona podría ser contrarrestado por la mayor seguridad de los rescatados en tal lugar, si hay en efecto una obligación legal de rescatar. Si la compensación es completa, habrá tantos rescatadores potenciales en la playa como habría si no existiese ninguna obligación.[43] Pero si no hay *más* rescatadores, no estará claro qué se ganará imponiendo una responsabilidad.

Puede parecer que la responsabilidad impondría un costo sólo a quien, si no fuese por la responsabilidad, *no* intentara un rescate; que los altruistas no se verían afectados. Pero aun el altruista desea tener una elección en el momento de la crisis acerca de intentar o no un rescate que podría ser peligroso para él; no desea verse coaccionado por la ley. Y uno de los beneficios que recibe una persona por ser un altruista es el del reconocimiento público, como lo sugiere la pequeña fracción de las caridades que son anónimas. La responsabilidad eliminaría este beneficio al hacer imposible que un rescatador demostrara que estaba motivado por el altruismo y no por las sanciones legales si no rescataba.

[42] Esta observación es congruente con el fallo dado en Ploof *vs*. Putnam, antes VI.4; porque es poco probable que el muelle no se haya construido debido a la preocupación de que el propietario pudiera ser demandado si su empleado impidiera que atracara un barco en apuros.

[43] Richard L. Hasen, "The Efficient Duty to Rescue", 15 *Int'l Rev. Law & Econ.,* 141 (1995).

Una alternativa para la responsabilidad y el altruismo como incentivos para el rescate es la restitución (antes IV.15). Sin embargo, el derecho común es más cauteloso acerca de compensar a la gente por brindar beneficios externos que acerca de hacerla pagar por los costos externos que impone.[44] Permitir que alguien que confiere beneficios externos externalice algunos de sus costos es una técnica más común, para tratar con los beneficios externos, que otorgarle un derecho de restitución; una persona que salva a una ciudad derribando una casa que se encuentra en la ruta de un incendio no tendrá que compensar al dueño de la casa; pero tampoco tendrá derecho a cobrar a los residentes de la ciudad por el beneficio que les ha conferido o siquiera por los costos en que haya incurrido. La diferencia radica en los costos de la administración de un sistema del derecho común. Las sanciones legales por imponer costos externos disuaden; especialmente cuando el estándar legal es la de la negligencia antes que el de la responsabilidad estricta, habrá pocos juicios legales. Pero un derecho de restitución crea un derecho legal siempre que se confiere un beneficio. Permitir que el benefactor externalice algunos de sus costos es un método más burdo, pero también más barato, para estimular la provisión de beneficios externos.[45]

Por último, adviértase que, en el caso Mouse (véase antes VI.4), el rescatador (el pasajero que arrojó por la borda el cofre de Mouse) estaba salvándose a sí mismo, así como a los demás, por lo que no necesitaba la motivación de una recompensa; ni tampoco lo habría disuadido la responsabilidad de intentar el rescate.

VI.10. La función de las indemnizaciones por cuasidelitos

El mantenimiento de la credibilidad del sistema de los cuasidelitos requiere que un demandado que ha sido declarado culpable pague daños y perjuicios por lo menos tan grandes como *L* en la Fórmula de Hand. Los daños deben pagarse a la víctima antes que al Estado porque, de otro modo, la víctima no tendrá ningún incentivo para demandar, y tal incentivo es esencial para el mantenimiento del sistema de los cuasidelitos como un disuasivo eficaz, creíble, de la negligencia. También podría tomar demasiadas precauciones. La segunda consiste en impedir que las víctimas tomen demasiadas precauciones. Recuérdese que, en el cuadro VI.1, el método óptimo de la prevención era que el ferrocarril gastara 50 dólares en un guardachispas y que el agricul-

[44] "Note, Efficient Land Use and the Internalization of Beneficial Spillovers: An Economic and Legal Analysis", 31 *Stan. L. Rev*., 457 (1979).

[45] Por lo que toca a la economía de la restitución, véanse William M. Landes y Richard A. Posner, "Salvors, Finders, Good Samaritans, and Other Rescuers: An Economic Study of Law and Altruism", 7 *J. Legal Stud.,* 83 (1978); Saul Levmore, "Obligation or Restitution for Best Efforts", 67 *S. Cal. L. Rev.,* 1411 (1994).

tor gastara 25 dólares en la reubicación de sus cultivos de lino más alejados de las vías, pero que si el ferrocarril no hacía nada el agricultor podría prevenir todavía el accidente, aunque a un costo de 110 dólares, desplazando sus cultivos mucho más atrás. Si el agricultor no fuese compensado por la negligencia del ferrocarril, éste seguiría adelante negligentemente, aunque fuera castigado por el Estado por dañar los cultivos del agricultor, sabiendo que el agricultor tendrá un incentivo para incurrir en un costo de 110 dólares a fin de evitar los daños esperados de 150 dólares y que, si lo hace, se evitará el accidente y el ferrocarril quedará liberado.

Evitar la sobrecompensación es tan importante como evitar la compensación insuficiente. La sobrecompensación puede aumentar el número de accidentes haciendo que las posibles víctimas sean descuidadas, y puede hacer que la atención médica sea más costosa si la responsabilidad por mala práctica es excesiva y los dispositivos de seguridad son más costosos porque la responsabilidad de los productos excesivos aumenta el precio de los dispositivos de seguridad por encima del costo esperado de defectos en los dispositivos.[46]

VI.11. Indemnizaciones por pérdida de la capacidad de ganar ingresos

Cuando un accidente incapacita a la víctima, el tribunal, en lugar de ordenar que el demandado haga pagos periódicos durante el periodo de la incapacidad (análogos a los pagos de pensión alimenticia al excónyuge), le ordenará que pague a la víctima en una sola exhibición calculada de tal modo que se iguale al valor presente de la corriente futura de los ingresos esperados que se han perdido. Esto difiere de la mera multiplicación de los ingresos de cada periodo de incapacidad por el número de periodos. Eso compensaría en exceso a la víctima (suponiendo que no hay inflación durante el periodo de la incapacidad, como veremos enseguida). Porque al final de la incapacidad habría recibido la víctima no sólo la suma de los pagos periódicos sino también los intereses de esa suma, los que no habría recibido si el pago se hubiera hecho periódicamente y no en una sola exhibición al principio. La exhibición única debe ser igual al precio que la víctima habría tenido que pagar para comprar una anualidad calculada con el objeto de rendir el pago periódico por la duración esperada de la incapacidad, y no más. Éste es el valor presente de la pérdida futura, y su pago en una sola exhibición es preferible a los pagos periódicos extendidos en el futuro. Así se economiza en los gastos administrativos y se evitan los efectos de desincentivo del hecho de ligar la recepción continua de dinero a la continuación de la incapacidad. Habiendo recibido la exhibición única, la víctima tiene todos los incentivos para superar su incapacidad más pronto de lo que se había pronosticado. En cambio,

[46] Véase Paul H. Rubin y Joanna M. Shepherd, "Tort Reform and Accidental Deaths", 50 *J. Law & Econ.*, 221 (2007).

un sistema de pagos periódicos por la incapacidad sería el equivalente a un impuesto de 100% sobre el ingreso ganado.[47] Sin embargo, un efecto potencial en sentido contrario es que el incentivo para exagerar las lesiones en el juicio es menor bajo el enfoque de los pagos periódicos. Éstos implican la revisión periódica de la incapacidad del demandante y, por lo tanto, una mucho mayor probabilidad de que se detecte cualquier exageración antes de que se complete el pago, lo que reduce el valor esperado de los pagos y en consecuencia el beneficio de la exageración. Pero la revisión periódica significaría más trabajo para los tribunales.

Los tribunales han tenido dificultades para determinar los daños y perjuicios en los casos que implican a las amas de casa incapacitadas. El enfoque más comúnmente empleado por los tribunales es la valuación de los servicios de un ama de casa sumando las cantidades que se requerirían para contratar proveedores de los diversos componentes de estos servicios (limpieza, cuidado de los hijos, cocina, etc.) —el enfoque del "costo de reposición"—, pero es poco satisfactorio porque olvida el costo de oportunidad. El valor mínimo de los servicios de un ama de casa, y por ende el costo para la familia si se eliminan esos servicios, son el precio que habría tenido el tiempo de la señora en el siguiente uso mejor. Supongamos que la señora había sido capacitada como abogada y podría haber ganado 100 000 dólares trabajando para una firma de abogados pero en lugar de eso escogió ser un ama de casa, y que los diversos servicios que ella realizaba como ama de casa podrían haber sido contratados en el mercado por 20 000 dólares. Dado que ella escogió permanecer en el hogar, presumiblemente sus servicios en el hogar eran valuados por la familia en por lo menos 100 000 dólares;[48] en caso contrario, la familia podría haber aumentado su ingreso real si ella trabajara como abogada y contratara a otros para que realizaran sus funciones hogareñas.[49] Por lo tanto, la pérdida causada cuando quedó incapacitada fue por lo menos de 100 000.

[47] El efecto impositivo se agravaría por el hecho de que el trabajo impone costos (impuesto al ingreso, traslado al trabajo, ropas de trabajo, etc.) que se evitan permaneciendo en casa y recibiendo pagos de incapacidad. Sin embargo, el trabajo podría generar un ingreso no pecuniario que sería igual o mayor que estos gastos… o podría generar una desutilidad adicional.

[48] Esto omite, a fin de simplificar la exposición, el incentivo que otorga a una mujer el sistema del impuesto al ingreso para que permanezca en el hogar aunque el valor de sus servicios fuera del hogar sería mayor que el de sus servicios como ama de casa. Véase después XVIII.8. Otra posibilidad es que la mujer obtenga un ingreso no pecuniario, quizá bajo la forma del ocio, por el hecho de permanecer en el hogar. Supongamos que el valor de los servicios de la esposa en el hogar asciende a 100 000 dólares y se compone de 60 000 dólares de los servicios prestados y 40 000 dólares del ocio producido. Si el valor de su tiempo ocioso no se ve afectado por la incapacidad, el costo de la incapacidad es de 60 000 dólares, antes que de 100 000 dólares al año. Pero también es posible que el trabajo fuera del hogar hubiera producido un ingreso no pecuniario por encima de su sueldo, en cuyo caso esa cifra podría subestimar el costo del accidente para ella.

[49] La decisión de que la señora permaneciera en el hogar podría haber sido muy racional, porque sus habilidades como ama de casa, particularmente en el cuidado de los hijos, podrían

Aunque podría haber sido mayor, así como el valor (los ingresos de toda la vida descontados) de una cantante de ópera podría superar a su valor en una ocupación alternativa, la valuación de los servicios de un ama de casa resulta difícil debido a la ausencia de un mercado explícito de amas de casa.[50] Sin embargo, los tribunales permiten los testimonios acerca de la calidad de los servicios hogareños del ama de casa. Éste es un método oblicuo para evitar el error de la valuación de tales servicios al costo de los trabajadores domésticos.[51]

El enfoque del costo de oportunidad tiene sus propios problemas de estimación. Por ejemplo, si una mujer se convierte en ama de casa (y permanece así), su capacidad para ganar ingresos en el mercado no alcanzará el nivel que habría alcanzado si no se hubiera convertido en ama de casa. Propiamente aplicado, el concepto del costo de oportunidad requeriría la estimación de lo que habrían sido sus probables ingresos en el mercado (restando cualquier costo de inversión, por ejemplo, en educación, relacionada con su empleo en el mercado) si hubiese entrado al mercado en el momento en que, por el contrario, se convirtió en un ama de casa.

Cuando los ingresos perdidos a causa de una lesión incapacitante se habrían acrecentado a lo largo de un periodo prolongado, tanto los supuestos formulados acerca de los cambios futuros de los ingresos de la víctima como acerca de la tasa del interés que se empleará para descontar tales ingresos a su valor presente pueden afectar grandemente el monto de la compensación. En el cuadro VI.2 se muestra el efecto del descuento. El punto de partida para

haber superado en gran medida a lo que la familia habría podido obtener en el mercado por el mismo precio. Véase antes V.3.

[50] Pero véase el estudio de Gronau citado antes, en V.3, nota 11, que trató de estimar el valor de la producción del ama de casa típica. Y adviértase la dependencia de la estimación de la pérdida respecto del hecho de que la trabajadora doméstica contratada no valga más que su sueldo. El ama de casa contratada puede producir un excedente del consumidor. Es decir, su producto total puede exceder a su salario (el producto marginal), si el mercado de trabajo es competitivo. Otra pregunta: ¿debieran las indemnizaciones del cuasidelito incluir las rentas económicas, es decir, el ingreso en exceso del costo de oportunidad? Sí, por dos razones. La primera es que la víctima de un cuasidelito a la que no se compensa en su totalidad tomará demasiadas precauciones. La segunda, que las rentas creadas por un recurso que es de por sí escaso (a diferencia de la escasez artificial provocada por un monopolista) reflejan el valor económico real de los beneficios otorgados por el recurso. Por tanto, la pérdida social engendrada por la destrucción de ese recurso se mide correctamente mediante la pérdida en las rentas.

[51] Supongamos que un profesor de la escuela de derecho que percibe un sueldo de 100 000 dólares queda incapacitado. Podría haber ganado 200 000 dólares ejerciendo la abogacía. ¿Cuáles son sus ingresos perdidos? Haciendo a un lado el efecto de los impuestos, los 100 000 dólares extra pueden ser sólo una compensación de las incomodidades de la práctica en comparación con la enseñanza; en tal caso, 100 000 dólares es una medida más real de lo que le ha costado la discapacidad. Sin embargo, si esto es correcto, implica que el abogado discapacitado que gana 200 000 dólares al año debería obtener mucho menos, porque su discapacidad elimina las incomodidades de la práctica. Pero, desde luego, puede que no haya ningún tipo de incomodidades para él.

CUADRO VI.2. *Valor presente de los ingresos futuros de 25 000 dólares al año, para varios periodos y tasas de descuento*

	Tasa de descuento			
Periodo	*2%*	*5%*	*10%*	*12%*
10 años	$224 565	$193 043	$153 615	$141 255
20 años	408 785	311 555	212 840	186 735
30 años	501 603	384 313	235 673	201 138

la investigación es el perfil salarial por edad para la ocupación de la víctima. Si la víctima era un camionero de 25 años de edad, necesitaríamos conocer los salarios no sólo de los camioneros de 25 años de edad sino también, suponiendo que la incapacidad es permanente y la jubilación ocurre a los 65 años, los salarios de los camioneros de 26 a 64 años de edad. El paso siguiente es la determinación de la manera como cambiarán probablemente los salarios de esa ocupación en el futuro. Muchos de los factores que podrían alterar el nivel salarial de una ocupación particular son muy difíciles de prever —como es el caso de los cambios en la demanda de la producción de la industria en la que el trabajador está empleado[52] o en el nivel de la sindicalización en la industria— y tendrán que omitirse generalmente; pero no todos. Las fuentes previsibles del cambio salarial incluyen la probabilidad del despido basada en la experiencia anterior del empleo en la industria, la creciente productividad laboral y la inflación. La productividad es la razón entre la producción y el insumo. Un aumento de la productividad del trabajo es un aumento de la cantidad de producción por hora de trabajo. Al reducir los costos de producción, un aumento de la productividad del trabajo le permite al empleador pagar salarios mayores (y la competencia por los trabajadores lo obligará a hacerlo). El aumento de la productividad del trabajo parece ser responsable de un aumento medio anual de cerca de 3% en los ingresos reales (es decir, ajustados por la inflación) de los trabajadores.[53]

Habiendo obtenido una estimación aproximada de los salarios reales que habría ganado el camionero en cada año de su vida de trabajo si no se hubiera incapacitado, el tribunal deberá multiplicar enseguida los salarios estimados para cada año por la probabilidad de que el camionero permanezca con vida al final de ese año. El tribunal también podría ajustar la estima-

[52] Si la demanda aumenta (disminuye), entonces la tasa salarial aumentará (disminuirá) si la curva de la oferta de mano de obra tiene una pendiente positiva. La pendiente positiva implica cierta escasez, de modo que el empleador tiene que pagar mucho más para apartar a los trabajadores de sus empleos existentes a fin de satisfacer su necesidad de mayores insumos de trabajo cuando su producción aumente.

[53] El efecto se hace sentir incluso en las industrias donde no está aumentando la productividad: los empleadores de tales industrias deben aumentar sus salarios hasta niveles competitivos con los de otras industrias, o perderán su fuerza de trabajo.

ción de sus salarios reales para reflejar el efecto sobre los salarios nominales de la inflación que se espera, aunque ello no es necesario, como veremos pronto.

Las expectativas de los inversionistas tal como se reflejan en las tasas de interés sobre los instrumentos financieros sin riesgo a largo plazo, como los bonos del gobierno de los Estados Unidos, se pueden utilizar para estimar la inflación a largo plazo. Una tasa de interés tiene tres componentes principales.[54] El primero es el costo de oportunidad del capital restando todo riesgo de pérdida y cualquier expectativa de inflación (o deflación). El segundo es el premio por el riesgo que se necesita para compensar al inversionista por la posibilidad de que jamás obtenga la devolución de su capital, un premio que se verá afectado por la actitud del inversionista hacia el riesgo.[55] Pero el riesgo no es un factor en los valores del gobierno de los Estados Unidos. El tercer componente es la tasa de inflación esperada durante el periodo en el que el préstamo estará vigente. Si el préstamo es por un año y se espera que el poder de compra del dólar disminuya en 4% durante ese año, entonces aunque no haya ningún riesgo de impago exigirá el prestamista el costo de oportunidad de quedarse sin su capital durante un año, más el 4% para compensarse por la disminución del poder de compra de ese capital durante el año. Una estimación razonable del costo real sin riesgo del capital es de 2%;[56] la tasa de interés actual de los bonos gubernamentales a largo plazo se acerca al 7%;[57] y la resta de la primera cifra menos la segunda se traduce en una tasa de inflación esperada a largo plazo de 5%, que debiéramos añadir entonces (compuesto) a la estimación de cada año de los ingresos perdidos por el trabajador.

El paso final es descontar nuestras estimaciones hasta el valor presente: una suma que, cuando se invierta hoy en día se acumulará hasta una suma

[54] Omitiendo los gastos del prestamista por concepto de negociación y administración del préstamo. *Cf.* xv.3.

[55] Supongamos que un préstamo de 100 dólares corre un riesgo de impago de 1% cada año. Si la tasa de interés sin el riesgo de impago es de 5%, un prestamista que sea neutral hacia el riesgo demandará 6%, porque $100 × 0.99 × 1.06 = $105. (Estas cifras están redondeadas.) Un prestamista que sienta aversión por el riesgo demandará más de 6%, mientras que un prestamista que prefiera el riesgo demandará menos.

[56] Véanse los estudios económicos citados en Doca *vs.* Marina Mercante Nicaragüense, S. A., 634 F.2d, 30, 39 n. 2 (2d Cir., 1980). Por cierto, Doca fue la primera decisión judicial que empleó la metodología económica moderna para determinar los ingresos perdidos en un caso de cuasidelito. Desde entonces ha habido muchas otras, como lo ilustran las siguientes: Jones & Laughlin Steel Corp. *vs.* Pfeifer, 462 U.S. 523 (1983); O'Shea *vs.* Riverway Towing Co., 677 F.2d, 1194 (7th Cir., 1982); Culver *vs.* Slater Boat Co., 688 F.2d, 280 (5th Cir., 1982) (en banc); In re Connecticut National Bank, 928 F.2d, 39 (2d Cir., 1991); Quiñones-Pachecho *vs.* American Airlines, Inc., 979 F.2d, 1 (1st Cir., 1992); Hollie *vs.* Korean Airlines Co., 834 F. Supp. 65 (S.D.N.Y., 1993). Véase en lo general Michael I. Krauss y Robert A. Levy, "Calculating Tort Damages for Lost Future Earnings: The Puzzles of Tax, Inflation and Risk", 31 *Gonzaga L. Rev.*, 325 (1996).

[57] Ésta es una tasa nominal; incluye la inflación esperada y el costo real del capital.

(principal más intereses) igual a los ingresos perdidos estimados para cada año que el demandante permanezca incapacitado para trabajar. Entre mayor sea la tasa de interés, menor será el valor presente (véase el cuadro VI.2), porque una tasa mayor hará que el valor crezca con mayor rapidez a lo largo del tiempo, de modo que entre mayor sea la tasa menor será la inversión presente necesaria para generar los ingresos anuales futuros que el demandante ha perdido. Supongamos que la tasa de interés de los bonos a largo plazo —es decir, una tasa de interés con un componente de inflación— se utilizara para este descuento. Ésta sería una tasa elevada y conduciría por lo tanto a un pago de daños menor que si se utilizara una tasa de interés más baja, pero el demandante no se vería perjudicado por este enfoque. La tasa sería elevada por contener un factor de inflación que ya se habría utilizado para aumentar la estimación de los ingresos futuros perdidos por el demandante. Si calculáramos éstos —en términos libres de la inflación—, cada estimación sería 5% (compuesto) menor, pero nosotros descontaríamos estas estimaciones a un valor presente utilizando una tasa de interés 5% menor. Como se trata de salir tablas, no es necesario en absoluto utilizar las estimaciones de la inflación en el cálculo de los daños del cuasidelito. Se pueden utilizar tanto estimaciones salariales como tasas de interés haciendo abstracción del fenómeno de la inflación.

Se considere o se ignore la inflación, no se debe utilizar la tasa de interés sin riesgo (la tasa de los valores emitidos o asegurados por la federación). La corriente de ingresos perdidos de la que el fallo de daños otorgado es el equivalente de valor presente no carece de riesgo: la muerte, el desempleo y la incapacidad por otras causas podrían haberla cortado, y otras contingencias podrían haberla aumentado (como trabajar después de la edad de jubilación normal). Por lo tanto, si el demandante siente aversión por el riesgo, considerará que el equivalente sin riesgo de su corriente riesgosa de ingresos perdidos vale más para él. Debe añadirse un factor de riesgo a la tasa de descuento para disminuir el valor presente hasta un nivel que confiera la misma utilidad que la corriente riesgosa de ingresos que se intenta remplazar.

Aunque las probabilidades de la muerte, el desempleo, etc., se tomaran en cuenta al estimar la corriente de ingresos perdidos, no se incurre en un conteo doble existencial rechazar, a causa de esas probabilidades, el uso de la tasa de interés sin riesgo para hacer el descuento de los ingresos futuros hasta su valor presente. Las probabilidades simplemente determinan el monto estimado de la pérdida futura. El hecho de que sea una pérdida incierta hace que revista más gravedad para la víctima (si siente aversión al riesgo) que si fuera cierta.

Una última complicación: La teoría de la producción hogareña discutida en el capítulo V implica que los ingresos reales de un individuo no se limitan al ingreso de mercado que gana en 40 de las 168 horas de una semana.[58] Un

[58] Véase Neil K. Komesar, "Toward a General Theory of Personal Injury Loss", 3 *J. Legal Stud.*, 457 (1974).

accidente grave que incapacita a una persona para trabajar en el mercado podría dañar también a la productividad de sus horas fuera del mercado, las que habría estado empleando para producir recreación, amor u otros bienes hogareños. Las tasas salariales de los segundos empleos (el dobleteo) podrían utilizarse para estimar los costos de oportunidad del hecho de estar incapacitado para el uso productivo de las horas no dedicadas a ganar ingresos en el mercado. Pero, si se hace esto, la porción de los salarios de una persona en segundos empleos que representa una compensación por los costos del trabajo (incluido el ingreso no pecuniario proveniente del ocio que se haya perdido, y cualesquiera riesgos o incomodidades implicados en el empleo) deberá restarse para determinar la pérdida neta del trabajador derivada de la incapacidad.

VI.12. Indemnizaciones por el dolor y el sufrimiento, el problema de la valuación de la vida humana y el riesgo de la compensación excesiva

La confusión del lego, destacada en el capítulo I de este libro, entre la pérdida pecuniaria y la pérdida económica se encuentra en la base de las críticas al otorgamiento de daños y perjuicios (con frecuencia muy sustanciales) por el dolor, la desfiguración, la pérdida de movilidad y otras formas del sufrimiento infligido por los accidentes, aun si los ingresos del trabajo no se ven afectados. Las críticas son poco sólidas. Tales pérdidas imponen costos de oportunidad. La gente pagará para evitarlas y demandará un pago por el riesgo de incurrir en ellas.

El otorgamiento de indemnizaciones por el dolor y el sufrimiento, aun cuando sea aparentemente generoso, podría compensar de forma deficiente a las víctimas gravemente incapacitadas por los accidentes. Dado que la pérdida de la visión o de los miembros del cuerpo reduce la cantidad del placer que puede comprarse con un dólar,[59] con frecuencia será necesaria una cantidad de dinero muy grande para colocar a la víctima en la misma posición de satisfacción relativa que tenía antes del accidente. El problema es más agudo en un caso de muerte. La mayoría de los individuos no cambiarían sus vidas por nada menor que una suma de dinero infinita si el intercambio debiera ocurrir de inmediato, ya que tendrían muy poco tiempo para disfrutar el beneficio de la venta. Pero no puede ser correcto que el otorgamiento apropiado de daños y perjuicios en un caso de muerte sea infinito. Además de la imposibilidad de tal compensación, esto implicaría que la tasa óptima de los accidentes fatales fuese cero o algo muy próximo a cero (en algún punto, los esfuerzos para promover la seguridad, por ejemplo, haciendo que las ambulancias no vayan a más de 8 km/h, están destinados a costar vidas), y es evi-

[59] Éste es un ejemplo de la manera como una lesión puede disminuir los ingresos no pecuniarios derivados de la actividad fuera del mercado. Véase antes VI.11.

dente que los individuos no están dispuestos, individual o colectivamente, a incurrir en los costos necesarios para reducir la tasa de los accidentes fatales tan drásticamente.

Los tribunales han resuelto el complejo problema de la valuación apropiada de la vida omitiéndolo en gran medida. La indemnización en un caso de muerte se limita generalmente a la compensación de la pérdida pecuniaria de los sobrevivientes (en algunos estados de la Unión Americana, la pérdida del caudal hereditario de la víctima fallecida), más los gastos médicos y cualquier dolor y sufrimiento experimentado por la víctima antes de morir. La pérdida pecuniaria de los sobrevivientes es la de los ingresos perdidos de la víctima menos sus gastos de vida. La prueba es así la misma que en los casos de incapacidad, excepto que, dado que los gastos personales de la víctima se eliminan por la muerte, pero no por la incapacidad, deberán sustraerse de la cantidad otorgada en un caso de muerte, pero no en un caso de incapacidad. ¡El supuesto implícito es que la persona que ha muerto no obtenía ninguna utilidad del hecho de estar viva![60]

Alguien que pierde con frecuencia cuando una persona muere o resulta incapacitada, además de sus sobrevivientes, es su empleador. Los empleadores invierten dinero en la capacitación de sus empleados y esperan recuperarse con la mayor productividad que la capacitación generará. El capital

[60] Sin embargo, hay un movimiento incipiente hacia el otorgamiento de daños por la utilidad perdida de vivir ("daños hedonísticos"). Véanse, por ejemplo, McGee *vs*. A C and S, Inc., 933 So. 2d 770, 775 (La., 2006); Golden Eagle Archery, Inc. *vs*. Jackson, 116 S.W.3d 757, 768 (Tex., 2003). Véase una discusión en Erin Ann O'Hara, "Hedonic Damages for Wrongful Death: Are Tortfeasors Getting Away with Murder?", 78 *Geo. L. J.,* 1687 (1990); Richard A. Posner, *Aging and Old Age,* 307-308 (1995); Sherwin Rosen, "The Value of Changes in Life Expectancy", 1 *J. Risk & Uncertainty,* 285 (1988). Sin embargo, la mayoría de los tribunales se niegan a otorgar los daños hedonísticos. Véanse Thomas R. Ireland, "The Last of Hedonic Damages: Nevada, New Mexico, and Running a Bluff", 16 *J. Legal Econ.*, 91 (2009); Eric A. Posner y Cass R. Sunstein, "Dollars and Death", 72 *U. Chi. L. Rev.*, 537, 539, 540, 542-543 (2005); Andrew Jay McClurg, "It's a Wonderful Life: The Case for Hedonic Damages in Wrongful Death Cases", 66 *Notre Dame L. Rev.*, 57 (1990). Victor E. Schwartz y Cary Silverman, en su artículo "Hedonic Damages: The Rapidly Bubbling Cauldron", 69 *Brooklyn L. Rev.*, 1037, 1039-1040 (2004), esbozan "los numerosos problemas con los daños hedonísticos. En primer lugar, contar la pérdida hedonística por separado y aparte del dolor y sufrimiento crea un riesgo significativo de doble indemnización. En segundo lugar, los daños hedonísticos amenazan a los límites legislativos y judiciales de la magnitud de los daños punitivos o no económicos, lo que puede conducir a la concesión de una ganancia inesperada. Por otra parte, las indemnizaciones por daños hedonísticos cuestionan los principios importantes de eficacia comprobada que subyacen a las disposiciones del homicidio culposo y a las acciones de supervivencia, que por lo general y sabiamente limitan la recuperación a la pérdida pecuniaria. La idea de los daños hedonísticos no sólo carece de una base conceptual, sino que la medida sufre de problemas de prueba. Debido a que los supuestos expertos no pueden ofrecer un estándar económico real para cuantificar el valor incalculable de la vida, el testimonio perjudica con frecuencia llevando a los miembros del jurado a una concesión arbitraria y exagerada. Por último, no existe una verificación sistémica confiable al final del proceso, ya que la naturaleza altamente subjetiva de los daños hedonísticos dificulta bastante la revisión significativa de la apelación".

humano así creado es un activo tan real del empleador como su maquinaria, y su destrucción es un costo igual de real. Algunos tribunales extranjeros han reconocido implícitamente esto al indemnizar al empleador por una lesión causada a un empleado.[61] Los tribunales del derecho común solían otorgar tales daños y perjuicios pero ya no lo hacen, considerándolos erróneamente como una implicación de que el empleador es "dueño" del empleado.

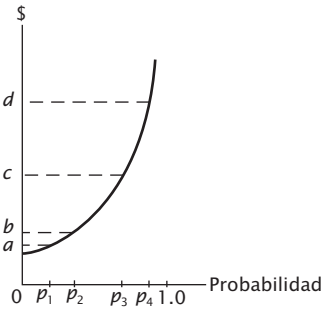
Un problema particularmente difícil es el de la valuación de la vida de un hijo. Aunque quizá no haya ninguna base para estimar los ingresos perdidos, una estimación mínima de la pérdida experimentada por los padres, que puede utilizarse como la base para el otorgamiento de daños a su favor, es su inversión de dinero y tiempo (este último monetizable sobre la base de los costos de oportunidad del mercado) en la crianza del hijo hasta la fecha de su muerte.[62] Los padres no habrían hecho la inversión si el valor esperado del hijo para ellos hubiese sido menor que el costo de la inversión. Por supuesto, el valor del hijo podría ser mucho mayor, pero la estimación del valor pleno presentaría graves dificultades, como en el caso de la estimación del pago de daños por la pérdida de los servicios de un ama de casa.

Este método de valuación se vuelve menos confiable a medida que el hijo crece, porque entonces los padres habrán recuperado más de su inversión bajo la forma de los servicios intangibles (por ejemplo, el hecho de ser brillante, bonito, tierno) que les prestó el hijo. Por otra parte, entre mayor sea el hijo, más fácil será pronosticar sus ingresos en el mercado.

El problema de la valuación de la vida de una persona para esta misma antes que para sus sobrevivientes podría ser solucionado, aunque sólo de manera parcial, distinguiendo entre los cambios de la utilidad *ex ante* y *ex post* generados por la actividad riesgosa. Si conduzco por la calle descuidadamente, crearé un riesgo de lesionar a muchas personas. Los economistas han calculado la cantidad de dinero que la persona media podría exigir por incurrir en un riesgo dado con base en estudios de las compensaciones salariales para los trabajadores de ocupaciones riesgosas, la respuesta de los precios de la vivienda ante la cercanía con sitios peligrosos, el uso de cinturones de seguridad, el hábito de fumar cigarrillos y otros comportamientos ante el riesgo de muerte, real o percibido. La división del "precio" que se cobra por asumir un riesgo determinado entre el riesgo da como resultado la estimación del valor de la vida. Por lo tanto, si el riesgo de muerte es 0.001 y el precio exigido para asumir el riesgo es de 7 000 dólares, el valor de la vida sería de 7 millones

[61] Football Club de Metz *c.* Wiroth, 1956 Recueil Dalloz 723 (Tribunal de Apelación de Colmar); Camerlo *c.* Dassary et Demeyere, 1958 Recueil Dalloz 253 (Tribunal de Apelación de Lyon). Sin embargo, es relevante que el empleado tuviera un contrato de duración fija con su empleador; si pudiera renunciar en cualquier momento se volvería incierto si el empleador hubiera recuperado o no su inversión en su capital humano. Véase XI.5 después.

[62] Wycko *vs.* Gnodtke, 105 N.W. 2d 118, 122 (Mich., 1960); Smith *vs*. City of Detroit, 202 N.W.2d 300 (Mich., 1972).



GRÁFICA VI.3

(7 000 ÷ 0.001)… que en realidad es la estimación promedio actual del valor promedio de la vida de un estadunidense.[63] Así, mientras que la valoración de la vida humana suena como una tarea ética o incluso metafísica, lo que realmente implica es la determinación del valor que la gente atribuye a la evitación de riesgos de muerte pequeños; la conversión en un "valor de la vida" se hace simplemente para determinar qué tan grande debe ser la concesión de la suma total del pago de daños en una exhibición para que sea equivalente. Por lo tanto, si la víctima potencial valúa en 7 000 dólares la eliminación de un riesgo de 0.001, cualquier precaución que elimine el riesgo a un costo menor a 7 000 sería eficiente. (En otras palabras, *PL* es de 7 000.) Puesto que *P* es de 0.001, *L*, que puede calcularse dividiendo 7 000 dólares (*PL*) entre 0.001 (P), debe ser de 7 millones. Si ésta es la cantidad que debe pagar el victimario que no toma precauciones que eliminen el riesgo a un costo menor a 7 000 dólares, si la muerte sobreviene como consecuencia de su omisión, será la cantidad correcta de disuasión.

Pero este enfoque no funcionará cuando la probabilidad de la muerte sea elevada. El hecho de que alguien pida sólo 7 000 dólares por incurrir en un riesgo de muerte de 0.001 no implica que pedirá sólo 1 169 000 dólares para jugar a la ruleta rusa[64] o 7 millones de dólares por incurrir en una muerte segura. Pero si inferimos de esto que el valor de la vida es siempre infinito, entonces *PL* será también infinito, sin que importe cuán pequeña sea *P*, y que los individuos no asumirán jamás ningún riesgo: una descripción obviamente falsa del comportamiento humano. Por lo tanto, parecería que el valor de la vida (el eje vertical en la gráfica VI.3) sube más deprisa que el riesgo de la muerte (el eje horizontal). Si todo incremento del riesgo produjera la misma desutilidad, la curva que relaciona la desutilidad con el riesgo sería una línea

[63] W. Kip Viscusi y Joseph E. Aldy, "The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World", 27 *J. Risk & Uncertainty*, 5, 6, 63 (2003).

[64] La probabilidad de muerte derivada de una ronda de ruleta rusa —0.167— equivale a una probabilidad de muerte de 0.001, como 1.169 millones de dólares son a 7 000.

recta. Su forma asintótica indica que los individuos pedirán mucho más dinero, por asumir un riesgo grande, que la suma computada multiplicando el dinero pedido por asumir un riesgo pequeño por el incremento del riesgo. Por lo tanto, *cd* en la gráfica VI.3 es mucho mayor que *ab*, aunque el incremento del riesgo compensado por *cd* ($p4 - p3$) sea el mismo que el compensado por *ab* ($p2 - p1$).

Incluso cuando la probabilidad de muerte es baja, se puede cuestionar el enfoque económico para calcular el valor de la vida. Recordemos las críticas por la atribución de probabilidades numéricas arbitrarias del capítulo I, como respuesta a la incertidumbre (riesgo no calculable). Supongamos que la observación del comportamiento de conducción de la gente muestra que no incurrirá en un costo de un dólar por la demora para llegar a su destino con el fin de reducir la probabilidad de morir en un accidente de tráfico por menos de 1 en 7 millones. De este modo, aplicando la fórmula del valor de la vida $v = c/p$, donde c es el costo de la evasión y p la probabilidad de la muerte si no se incurre en el costo, el valor de la vida (v) será de 7 millones de dólares ($1 \div 1/7\,000\,000 = 7\,000\,000$). Pero es probable que un conductor no tenga idea de qué es p. Por consiguiente, la "preferencia" que "revela" a través de su comportamiento al incurrir o no en c —siendo la preferencia revelada el método favorito de los economistas para determinar la valuación de la gente— es un índice de sus preferencias poco fiable.

Una complicación adicional es que algunas víctimas de accidentes obtienen escasa o ninguna utilidad de un otorgamiento de pago de daños. La víctima que muere, o que se sume en un coma permanente, no obtiene ninguna utilidad; la víctima que queda convertida en un parapléjico podría obtener muy poca utilidad. En estos y muchos otros casos, el accidente disminuye la utilidad marginal del ingreso para la víctima. (Esto supone, desde luego, que el otorgamiento de daños cubrirá sus gastos médicos, los cuales pueden ser enormes.) Un individuo racional desea obtener lo más posible por sus centavos y, por lo tanto, desearía reasignar el ingreso del periodo posterior al accidente, cuando su utilidad marginal del ingreso es baja, al periodo anterior al accidente, cuando es elevada.[65] Esto podría hacerse mediante un gravamen fiscal de los otorgamientos de daños en casos de muerte o de grave incapacidad permanente, ya que la recaudación del impuesto aumentaría (quizás haciendo posibles las reducciones de otros impuestos) el ingreso disponible de la población en general, la mayoría de cuyos miembros no están por supuesto incapacitados. Además, las víctimas potenciales de los accidentes podrían tener más cuidado porque su ingreso anterior al accidente sería más alto en promedio, por lo que tendrían más que perder a resultas de un accidente mortal o incapacitante.

[65] David Friedman, "What Is 'Fair Compensation' for Death or Injury?", 2 *Int'l Rev. Law & Econ*., 81 (1982). Una prueba de esta preferencia es que los individuos compran raras veces seguros contra el dolor y el sufrimiento. Pero ¿se le ocurre otra razón para que procedan así?

Pero esta discusión supone, quizás injustificadamente,[66] que el punto de partida apropiado para la determinación de la utilidad marginal del ingreso es la situación existente antes del accidente. Si se emplea por el contrario la perspectiva del yo futuro incapacitado contingente, se vuelve evidente que, a menos que el yo futuro esté destinado a ser irreversiblemente comatoso, bien podría asignar una utilidad marginal del ingreso mayor en el estado de incapacidad que el que asignaría el yo antes del accidente, y, por lo tanto, asignaría un valor menor a los gastos en consumo del yo anterior al accidente. Según esta visión, el derecho aplicable a los cuasidelitos opera como una especie de fiduciario para el yo incapacitado.

VI.13. La regla de los beneficios colaterales (fuente colateral)

Si una póliza de seguro contra accidentes me faculta para recibir 10 000 dólares por cierta clase de lesión accidental y yo sufro esa lesión en un accidente en el que el victimario es negligente, podré reclamar los 10 000 dólares de la compañía de seguros y obtener del victimario el pago pleno de los daños (que, supongamos, son 10 000 dólares), a condición de que yo no haya convenido en ceder mis derechos por el cuasidelito al asegurador (subrogación). Si se permite que el demandado esgrima mi póliza de seguro como una excepción contra la acción, se generaría una disuasión insuficiente. El costo económico del accidente, comoquiera que se pague, es 10 000 dólares, y si el fallo contra el demandado es cero disminuirá su incentivo para gastar hasta 10 000 dólares (descontados por la probabilidad de la ocurrencia) a fin de prevenir un accidente similar en el futuro. Menos obviamente, la doble recuperación no es una ganancia súbita para mí. Yo compré la póliza de seguro a un precio presumiblemente igual al costo esperado de mi lesión más el costo de adquisición de la póliza. La compañía podría, si quisiera, haberme exceptuado de la cobertura de accidentes en los que el victimario era responsable ante mí por el costo del perjuicio, o podría haberme exigido que le cediera cualesquiera derechos legales que pudiera yo tener a resultas de un accidente (esto es común, y se conoce como subrogación). En ambos casos, mi prima habría sido menor.

Algunos tribunales han tenido problemas cuando el beneficio colateral no derivaba del contrato sino que era "gratuito". Sin embargo, la mayoría de los beneficios gratuitos resultan ser aquellos por los que el beneficiario ha pagado indirectamente. Si un empleador otorga tratamiento médico sin car-

[66] Véanse Steven P. Croley y Jon D. Hanson, "The NonPecuniary Costs of Accidents: Pain-and-Suffering Damages in Tort Law", 108 *Harv. L. Rev.*, 1785, 1822-1827 (1995); Richard A. Posner, *Aging and Old Age*, 84-85 (1995); Richard A. Posner, "Are We One Self or Multiple Selves? Implications for Law and Public Policy", 3 *Legal Theory*, 23 (1997); y la discusión anterior del descuento hiperbólico en I.4.

go a sus empleados lesionados, esto sólo significa que el empleador paga su trabajo parte en dinero y parte en especie, de modo que el salario monetario sería mayor si los beneficios "gratuitos" fuesen menores.

Supongamos que un trabajador es despedido de su empleo porque alguien lo calumnia con su empleador; el trabajador demanda al calumniador y recupera daños y perjuicios por el periodo en el que estuvo desempleado a resultas del despido. ¿Debieran deducirse del fallo cualesquiera beneficios de desempleo que hubiera recibido? Como una primera aproximación, no debieran deducirse. El costo social de su desempleo es el valor de la producción que el trabajador habría realizado si hubiese conservado su empleo, y ese costo no disminuye por el hecho de que las consecuencias financieras para él se atemperaran por el seguro de desempleo. En la medida en que el gobierno financie el seguro de desempleo, no obstante, hay un argumento en favor de la deducción de los beneficios del pago de daños obtenido por el empleado, y de que se permita que el gobierno demande para recuperarlos. En realidad, el gobierno es otra víctima del accidente.

VI.14. Negligencia con seguro de responsabilidad y seguro contra accidentes, compensación por accidente automovilístico sin falta y limitación de las demandas por mala práctica médica

El sistema de la negligencia ha sido objeto de críticas adversas en los medios durante muchos años.[67] La crítica principal es que se trata de un sistema de compensación caro e inadecuado. La atención se centró inicialmente en los casos de accidentes automovilísticos, el tipo de negligencia más frecuente, y comenzaremos aquí, aunque en la actualidad existe una mayor preocupación por la mala práctica médica y la responsabilidad por los productos.

Algunos estudios revelan que los costos administrativos, principalmente los gastos legales, constituyen una fracción elevada de las cantidades totales pagadas a las víctimas de accidentes automovilísticos en los arreglos y los juicios, y que muchas personas lesionadas en dichos percances reciben poca o ninguna compensación, a veces porque la víctima misma fue negligente, otras veces porque el demandado no estaba asegurado y era insolvente o el conductor era un desconocido del tipo pega y huye.

Si la compensación es el único propósito del sistema de la negligencia, será en efecto un sistema defectuoso, ya que es costoso e incompleto. Pero su función económica principal, dada la existencia del seguro de accidentes y de responsabilidad, no es la compensación sino la disuasión de los accidentes ineficientes. Si el sistema genera ahorros sustanciales en los costos de los

[67] Véase en Guido Calabresi, *The Costs of Accidents: A Legal and Economic Analysis* (1970), un ataque vigoroso aunque ya superado con el tiempo, primordialmente económico, contra el sistema de la negligencia.

accidentes, sus pesados costos de administración, que se relacionan principalmente con la determinación de la responsabilidad, podrían estar justificados.

El seguro contra la responsabilidad automovilística es ahora casi universal, aunque esto se debe en parte al hecho de que los estados de la Unión Americana requieren que los conductores compren seguros de responsabilidad o presenten una prueba equivalente de responsabilidad financiera por los accidentes. Pero, aun sin compulsión, el seguro contra la responsabilidad sería común. El predominio de la aversión hacia el riesgo es una condición necesaria pero, sorprendentemente, no suficiente para este fenómeno. En teoría, hay un procedimiento más barato para evitar el riesgo de ser declarado responsable de un perjuicio automovilístico: no ser negligente. Con todo, en virtud de que los tribunales cometen errores y de que la negligencia contiene, como vimos, un componente de responsabilidad estricta, hay siempre un riesgo de ser juzgado negligente y, por ende, una demanda del seguro contra la responsabilidad.

Con ese seguro, sin embargo, el costo de un accidente para el victimario negligente ya no es la pérdida de la víctima; es el valor presente de cualquier aumento de la prima que el victimario pueda experimentar a resultas de ser declarado negligente. Debido a los costos de información, la hostilidad de los reguladores hacia las estructuras "discriminatorias" de las tasas de primas, y a las combinaciones de los riesgos asignados por mandato gubernamental que permiten aun a los conductores más peligrosos comprar seguros contra la responsabilidad a tasas sólo ligeramente mayores que las normales, las primas de los seguros contra la responsabilidad no se ajustan con precisión a los costos esperados de los accidentes de conductores particulares. Aunque las primas no son uniformes, las diferencias reflejan con frecuencia criterios, como el verse implicado en un accidente (independientemente de que el asegurado haya sido negligente o no) o el grupo de edad al que pertenezca el asegurado, que se relacionan sólo vagamente con la negligencia. Hay una varianza elevada en los costos esperados de los accidentes dentro de las clases de prima, de modo que el método de cálculo de las tasas de los seguros contra la responsabilidad disuade en exceso a algunos conductores y deficientemente a otros.

En un mercado de seguros no regulado, las compañías de seguros podrían cobrar a sus clientes primas relacionadas más estrechamente con la responsabilidad esperada.[68] Sin embargo, los errores seguirían siendo frecuentes debido al costo de la obtención de información suficiente acerca del

[68] O quizá no pudieran hacerlo, dependiendo del costo de hacer la diferenciación más fina. *Cf.* Frank A. Sloan, "Experience Rating: Does It Make Sense for Medical Malpractice Insurance?", 80 *Am. Econ. Rev. Papers & Proceedings,* 128 (mayo de 1990). Entre más accidentes se concentren en una fracción pequeña de los asegurados, más viable será la "tarifa de la experiencia" (el término general para el intento de subdividir al grupo de asegurados de acuerdo con las estimaciones de la probabilidad de una lesión que genere responsabilidad). El costo de la infor-

solicitante de un seguro para hacer una estimación precisa de la probabilidad de que cause perjuicios por los que un tribunal lo declararía responsable. De modo que habría más accidentes que si se prohibiera el aseguramiento de la responsabilidad. Pero, si las víctimas son plenamente compensadas, el seguro contra la responsabilidad es eficiente aunque se generen más accidentes. El asegurador y el asegurado pasan a una posición mejor, o el contrato de seguros no se habría celebrado, y nadie pasa a una posición peor bajo el supuesto de que se compensa plenamente a las víctimas de la negligencia del asegurado. Por lo tanto, aun cuando el seguro contra la responsabilidad sí mina un poco el poder disuasivo del derecho aplicable a los cuasidelitos, esto no lo convierte necesariamente en un sistema ineficiente para el control de los accidentes… pero sólo si la compensación del cuasidelito para todos los perjuicios es una compensación plena; y, para los perjuicios graves (especialmente los mortales), no lo es.

Una crítica permanente del sistema de la negligencia es que no es psicológicamente realista esperar que quienes no se disuaden de observar una conducta descuidada por temor a la lesión corporal sí lo hagan por temor a una condena monetaria; o bien, en el caso en que la negligencia de la víctima es un obstáculo para la recuperación, por la incapacidad para obtener del victimario una compensación. La crítica tiene mérito, pero requiere de matización. El argumento es inaplicable a los victimarios que no están ellos mismos en peligro personal, a los empleadores de los victimarios (tales como una compañía de camiones o de taxis) y a los accidentes donde el único peligro significativo es para la propiedad. Y no toma en cuenta el efecto de las tasas de las primas del seguro de responsabilidad que son tan elevadas, en reflejo de la gran responsabilidad esperada de algunos asegurados, que desalientan a algunas personas del proyecto de conducir automóviles. Por ejemplo, las primas elevadas que se cobran a los jóvenes conductores varones demoran el inicio de su conducción.[69] Cuanto menos se conduzca, habrá menos accidentes de conducción. Además, la experiencia con la compensación por el accidente automovilístico sin falta ha generado considerables pruebas empíricas, discutidas más adelante, de que la responsabilidad por los cuasidelitos sí disuade. Y, por último, aunque ésta no es una observación acerca de los accidentes automovilísticos, no debe olvidarse que la negligencia tiene un dominio mucho más amplio. Pocos comentaristas creen que la responsabilidad por la mala práctica médica o los productos no tiene ningún efecto sobre el comportamiento de los médicos y los fabricantes, respectivamente.[70] Algu-

mación es menor, y los beneficios de la tarifa de la experiencia en términos de la reducción de accidentes son mayores.

[69] Véase Michelle J. White, "The 'Arms Race' on American Roads: The Effect of Sport Utility Vehicles and Pickup Trucks on Traffic Safety", 47 *J. Law & Econ.*, 333 (2004).

[70] Véanse Theodore Eisenberg, "The Empirical Effects of Tort Reform" (Cornell Legal Studies Research Paper, 12-26, 2012), próxima publicación en *Research Handbook on the Econo-*

nos consideran que ha tenido un efecto demasiado grande, sobre todo, como veremos, en el comportamiento de los médicos.

En suma, "el derecho aplicable a los cuasidelitos provee algo importante en materia de disuasión"…[71] a pesar de lo cual algunos estados de la Unión Americana han promulgado leyes de compensación del accidente automovilístico sin falta, leyes que no se ocupan de la creación de mejores incentivos para evitar los accidentes, sino de tratar de aumentar la cobertura del sistema *y* disminuir el costo del seguro. Estas metas son incongruentes entre sí y con la meta de la disminución del número de accidentes.

El plan de Keeton-O'Connell,[72] que es el modelo para estos estatutos, ilustra el dilema. Según este plan, todo motorista debe tener una protección básica que lo faculte, en el caso de un accidente, para recuperar sus gastos médicos más los ingresos perdidos, independientemente de la negligencia del victimario o de su propia libertad ante la negligencia. El dolor y el sufrimiento no se compensan, y cualesquiera beneficios colaterales se deducen. La víctima podría renunciar a la protección básica y demandar en un juicio de cuasidelito, en la forma habitual, si sufre daños mayores de 10 000 dólares aparte del dolor y el sufrimiento. La protección básica es un seguro de la primera parte (accidente) antes que de la tercera parte (responsabilidad). El automovilista paga primas a su propio asegurador y recibe de éste el pago de daños y perjuicios. El victimario y su compañía aseguradora son responsables sólo si la víctima renuncia a la protección básica y demanda por el cuasidelito.

¿Por qué se excluye la indemnización por el dolor y el sufrimiento y se requiere la deducción de los beneficios colaterales? Aparentemente no porque los autores no consideren que el dolor y el sufrimiento son pérdidas reales, o porque consideren que los beneficios colaterales son meras ganancias súbitas, ya que no excluyen estos rubros en los accidentes graves. Pero necesitaban algún procedimiento para disminuir el pago promedio de daños a fin de evitar que el plan aumentara el costo del seguro. En virtud de que el plan compensa a las víctimas de conductores sin falta y a las propias víctimas con

mics of Torts (Jennifer H. Arlen [comp.], 2013); Darius N. Lakdawalla y Seth A. Seabury, "The Welfare Effects of Medical Malpractice Liability", 32 *Int'l Rev. Law & Econ.*, 356 (2012); Henry S. Farber y Michelle J. White, "Medical Malpractice: An Empirical Examination of the Litigation Process", 22 *RAND J. Econ.*, 199 (1991); W. Kip Viscusi, "The Performance of Liability Insurance in States with Different Products-Liability Statutes", 19 *J. Legal Stud.*, 809 (1990). Véase también David W. Prince y Paul H. Rubin, "The Effects of Product Liability Litigation on the Value of Firms", 4 *Am. Law & Econ. Rev.*, 44 (2002).

[71] Gary T. Schwartz, "Reality in the Economic Analysis of Tort Law: Does Tort Law Really Deter?", 42 *UCLA L. Rev.*, 377, 443 (1994). Véase también, con referencia específica a los decesos por el tráfico, Frank A. Sloan, Bridget A. Reilly y Christoph M. Schenzler, "Tort Liability *versus* Other Approaches for Deterring Careless Driving", 14 *Int'l Rev. Law & Econ*., 53 (1994) y, más adelante, las referencias de la nota al pie número 74.

[72] Robert E. Keeton y Jeffrey O'Connell, *Basic Protection for the Traffic Victim: A Blueprint for Reforming Automobile Insurance* (1965).

falta, su cobertura es mayor que la del sistema de los cuasidelitos. Por tanto, si la reclamación media no fuese menor bajo el plan que bajo el sistema de cuasidelitos existente, la cantidad total pagada por reclamaciones y, por ende, los costos de las primas de seguros serían probablemente mayores (aun suponiendo costos de administración menores) que bajo el sistema con falta. Los ahorros obtenidos de la deducción de los beneficios colaterales bien podrían resultar transitorios; los individuos podrían disminuir su seguro actual contra accidentes a fin de contrarrestar el seguro contra accidentes que se ven obligados a comprar bajo la protección básica. Sin embargo, la estrategia del plan es clara: aumentar el número de las víctimas de accidentes compensadas, pero disminuir la compensación media.

Un estudio ha descubierto que los estados de la Unión Americana cuyas leyes de no falta imponen restricciones severas a la responsabilidad de los cuasidelitos pueden esperar de 10 a 15% más de decesos por accidentes automovilísticos.[73] Este resultado puede parecer extraño, ya que los estatutos de no falta dejan intacta la responsabilidad del cuasidelito en los casos de decesos. No obstante, téngase presente el carácter probabilístico del cuidado: si la no falta induce una conducción más descuidada, habrá más accidentes y una parte de ellos serán fatales. Pero la no falta podría llevar a los conductores a manejar a la defensiva porque esperan que los otros conductores sean descuidados y, por lo tanto, podría reducir la tasa de accidentes o al menos mantenerla invariable.

En los últimos años, las primas del seguro de responsabilidad por mala práctica médica se han disparado, y los médicos están protestando enérgicamente y solicitando que se limiten los juicios por mala práctica. La gente a la que le preocupa el aumento continuo de los costos totales del cuidado de la salud en los Estados Unidos apoya sus protestas, a pesar de que las primas por mala práctica sólo representan alrededor de 1% de dichos costos totales. Sin embargo, existen pruebas de que la responsabilidad por mala práctica aumenta la calidad de la atención a la salud.[74] ¡Los opositores a la responsa-

[73] Elisabeth M. Landes, "Insurance, Liability, and Accidents: A Theoretical and Empirical Investigation of the Effect of No-Fault on Accidents", 25 *J. Law & Econ.*, 49 (1982). La investigación subsecuente ha confirmado que las leyes de no falta aumentan la tasa de decesos de los accidentes automovilísticos. Véanse Christopher J. Bruce, "The Deterrent Effect of Automobile Insurance and Tort Law: A Survey of the Empirical Literature", 6 *Law & Policy*, 67 (1984); J. David Cummins, Richard D. Phillips y Mary A. Weiss, "The Incentive Effects of No-Fault Automobile Insurance", 44 *J. Law & Econ.*, 427 (2001); Yu-Ping Liao y Michelle J. White, "No-Fault for Motor Vehicles: An Economic Analysis", 4 *Am. Law & Econ. Rev.*, 258 (2002); Adam Cohen y Rajeef Dehejia, "The Effect of Automobile Insurance and Accident Liability Laws on Traffic Fatalities", 47 *J. Law & Econ.*, 357 (2004).

[74] Patricia M. Danzon, "Liability for Medical Malpractice", en *Handbook of Health Economics,* 1339 (A. J. Culyer y J. P. Newhouse [comps.], 2000); Charles Silver y David A. Hyman, "The Poor State of Health Care Quality in the U.S.: Is Malpractice Liability Part of the Problem or Part of the Solution?" (Bepress Legal Series Working Paper, 222, 2004); Farber y White, nota 70.

bilidad reconocen esto, pero argumentan que aumenta demasiado la calidad! Dado que gran parte de la atención médica se paga a través del seguro antes que directamente por el paciente, los médicos podrían involucrarse en la medicina preventiva ("defensiva") más de lo óptimo con el fin de protegerse contra la posible responsabilidad por mala práctica,[75] ya que es probable que el paciente no asuma el costo. Otra preocupación es que, debido a que el seguro de mala práctica no se basa en la responsabilidad calificada por la experiencia —las primas no se basan en la experiencia de la responsabilidad personal de un médico, sino más bien en su especialidad (ginecología, cirugía ortopédica, etc.)—, sólo puede tener un efecto burdo sobre los métodos de los médicos o sobre el cuidado, salvo en la medida en que sus aseguradores presionen a los médicos a cambiar sus métodos a fin de reducir la cantidad de juicios por mala práctica. Por otra parte, en la práctica (y esto es cierto en general cuando el demandado en una demanda por un cuasidelito es una persona antes que una empresa u otra institución), el pago de daños obtenido (por lo general en un arreglo) en una demanda contra un médico se limita al tope de su cobertura de seguro; su riqueza personal no está en riesgo. Y puesto que es poco probable que los demandantes se nieguen a llegar a un acuerdo en el límite del seguro del demandado, "los médicos suelen establecer los límites de su póliza por debajo de las pérdidas esperadas de sus pacientes".[76]

Adviértase que, en la medida en que la responsabilidad por mala práctica transfiere la riqueza de los médicos a (algunos) pacientes, los costos totales de la atención a la salud no se ven afectados. El costo real de la responsabilidad por mala práctica (aparte del costo de la medicina defensiva) es el costo de los recursos valiosos consumidos por dicha responsabilidad, principalmente el tiempo de los abogados y de los peritos. La parte de la prima por mala práctica (aproximadamente la mitad) que paga por los abogados y los peritos debe entenderse como el costo de mantener un sistema para aumentar la seguridad médica y debe compararse con el beneficio de la reducción de la mala práctica.

Algunos críticos del cuasidelito de mala práctica médica proponen su abolición a favor de un régimen contractual. El nivel de atención se negociaría entre el médico y el paciente.[77] Dado que sólo hay dos partes en un contrato para el tratamiento médico, los costos de transacción no deben impedir una determinación negociada del nivel de cuidado óptimo, dadas tales consi-

[75] Véase Daniel Kessler y Mark McClellan, "Do Doctors Practice Defensive Medicine?", 111 *Q. J. Econ.,* 353 (1996). Kessler y McClellan descubrieron un efecto sustancial; otros estudios descubrieron uno mucho menor o encontraron que la medicina defensiva mejora el bienestar del paciente de manera efectiva. Véase, por ejemplo, Praveen Dhankhar, M. Mahmud Khan y Shalini Bagga, "Effect of Medical Malpractice on Resource Use and Mortality of AMI Patients", 4 *J. Empirical Legal Stud.*, 163 (2007).

[76] Jennifer Arlen, "Contracting over Liability: Medical Malpractice and the Cost of Choice", 158 *U. Pa. L. Rev.*, 957, 987 (2010).

[77] Véanse Richard A. Epstein, "Medical Malpractice: The Case for Contract", 1 *Am. Bar Foundation Res. J.*, 87 (1976), y otros trabajos citados en Arlen, nota 76 anterior, en 960 n.5.

deraciones como el precio y el miedo al error médico de parte del paciente… o eso podría pensar uno. Pero, aparte del hecho de que muchos tratamientos médicos involucran a múltiples partes al final del tratamiento (por ejemplo, internista, cirujano, radiólogo, anestesiólogo, enfermeras, técnicos, hospital), elevando los costos de transacción, existe un problema de acción colectiva. Al saber que da igual si demanda o renuncia a la responsabilidad de parte del médico, pues ello tendrá poco efecto sobre la probabilidad de un error médico —es poco probable que un error de esta índole sea deliberado, y un médico no tendrá menos cuidado de manera deliberada porque un paciente haya renunciado a su derecho de demandar por mala práctica médica—, el paciente tendrá pocos incentivos para resistir la renuncia. (De hecho, se beneficiará gratuitamente de la negativa a renunciar de otros pacientes y de esta manera motivará al médico para que sea más cuidadoso.) También hay un problema de selección adversa: los pacientes sanos o intrépidos renunciarán a la responsabilidad contractual por error médico, y los enfermos o temerosos insistirán en ella. Al no saber exactamente la razón por la que un paciente quiere la responsabilidad, el médico puede cobrar honorarios que, en promedio, serán mayores que el costo real mayor del error esperado. Dado que muchos pacientes no buscarán protección contractual contra la mala práctica, la alternativa contractual a la responsabilidad del cuasidelito tendrá menos efecto en incentivar a los médicos a evitar la mala práctica que la responsabilidad del cuasidelito, que rige a todos los médicos.[78]

VI.15. Los cuasidelitos intencionales

Hasta ahora, este capítulo se ha ocupado de los cuasidelitos accidentales o, como se llaman con frecuencia en la ley, no intencionales. Ahora debemos considerar la otra gran categoría del derecho aplicable a los cuasidelitos: los cuasidelitos intencionales. Como ocurre con muchas distinciones legales, ésta no es precisa, en parte porque el término "intencional" es vago. La mayoría de las lesiones accidentales son intencionales en el sentido de que el victimario sabía que podría haber disminuido la probabilidad del accidente tomando precauciones adicionales. El elemento de la intención es indiscutible cuando el cuasidelincuente es una empresa que puede prever a partir de la experiencia del pasado que infligirá cierto número de lesiones accidentales cada año. En cambio, en muchos cuasidelitos intencionales se encuentra muy atenuado el elemento de la intención, como ocurre cuando un cirujano que sin saberlo excede los límites del consentimiento expreso o implícito del paciente para un procedimiento quirúrgico es acusado de haber cometido una agresión. En

[78] El artículo de Arlen, nota 76 anterior, discute estas y otras objeciones económicas al régimen contractual para la responsabilidad médica.

el caso habitual de la agresión médica, se trata de saber si hubo una emergencia suficiente para justificar un procedimiento para el que no se había obtenido previamente el consentimiento del paciente. Esto depende a su vez de que los costos de la demora (tales como el riesgo de que pudiera deteriorarse la condición del paciente, y el peligro adicional de someterlo de nuevo a una anestesia general) excedan al valor para el paciente de una oportunidad de considerar si aceptará el procedimiento; en caso afirmativo, se encontrará el consentimiento implicado para el procedimiento. Este caso se asemeja mucho a los casos del contrato implicado discutidos al final del capítulo IV.

O consideremos los casos de las viejas escopetas de resorte. El demandado en Bird *vs.* Holbrook[79] era propietario de un valioso jardín de tulipanes situado aproximadamente a un kilómetro y medio de su casa. Aunque el jardín estaba amurallado, algunos tulipanes habían sido robados, de modo que el demandado preparó una carabina de resorte. La pava de un vecino se escapó y se introdujo al jardín. Un joven que perseguía a la pava se metió también al jardín tratando de recapturarla para su propietario, se tropezó con la carabina y resultó lesionado. El tribunal falló que el propietario era responsable de la lesión porque no había colocado avisos de que se había instalado una carabina de resorte. El incidente había ocurrido durante el día.

La controversia en este caso, tal como la consideraría un economista, era el acomodo apropiado de dos actividades legítimas: el cultivo de tulipanes y la crianza de pavos. El demandado tenía una inversión sustancial en el jardín de tulipanes; vivía a cierta distancia; y el muro no había resultado eficaz contra los ladrones. En una época de protección policiaca insignificante, una carabina de resorte podría haber sido el medio más eficaz, en términos de sus costos, para la protección de los tulipanes. Pero, dado que las carabinas de resorte no discriminan entre el ladrón y el allanador inocente, disuaden a los propietarios de animales domésticos de perseguir a sus animales hasta la propiedad de otra persona, aumentando así los costos (los costos del cercamiento o las pérdidas por los animales que se escapan) de la crianza de animales. El tribunal en el caso de Bird implicó un acomodo ingenioso: quien instale una carabina de resorte deberá colocar avisos en tal sentido. Entonces los propietarios de animales no se mostrarán reacios a perseguirlos hasta la propiedad que no tenga tales avisos. Un aviso no servirá de nada de noche, pero es más probable que los animales estén seguros entonces, y en todo caso pocos propietarios perseguirían a sus animales al oscurecer. El análisis resulta así ser el mismo que en un caso de negligencia: el caso arquetípico del cuasidelito no intencional.

El conjunto de cuasidelitos intencionales que son económicamente diferentes de los cuasidelitos no intencionales incluye el allanamiento deliberado (véase antes III.6), el asalto,[80] el ataque simple (por ejemplo, el atraco, por oposi-

[79] 4 Bing. 628, 130 Eng. Rep., 911 (C. P., 1828).

[80] Hacer un gesto amenazante que haga temer a una persona un ataque inminente. ¿Cuál in-

ción a la agresión médica o a la agresión técnica que se comete en una pelea de box ilegal), el fraude y la apropiación ilícita (la contrapartida cuasidelictiva del robo): cuasidelitos que se asemejan a delitos del derecho común tales como la violación, el homicidio, el robo simple, el robo a una casa y el latrocinio. Estos cuasidelitos y la correspondiente lista de delitos no implican un conflicto entre actividades legítimas (productivas) sino una transferencia coercitiva de riqueza al demandado en un contexto de costos de transacción bajos. Cuando los costos de transacción del mercado son bajos, debiera exigirse, como sabemos, a la gente que use el mercado si puede hacerlo y desista de su propósito si no puede hacerlo. El robo (o la apropiación ilícita) no es sólo un pago de transferencia sin costo —de modo que la objeción a su ejecución deba buscarse fuera de la economía—, aun cuando el ladrón valúe el bien robado tanto como su propietario.[81] Porque, si se permitieran tales transacciones, los propietarios gastarían mucho en la protección y los ladrones gastarían mucho para frustrar los esfuerzos de los propietarios. Sus gastos conjuntos serían un desperdicio desde un punto de vista social; este desperdicio es la objeción económica al robo.

La Fórmula de Hand es útil en tanto que ilustra la diferencia existente entre los cuasidelitos intencionales y los que no lo son, una diferenciación relevante en el derecho aplicable a los cuasidelitos. Consideremos el caso en el que un ferrocarril, gracias a que opera muchos trenes cada año, sabe con una confianza cercana a la certeza que matará a 20 personas al año en los cruces ferroviarios. ¿Es por ello un cuasidelincuente intencional? No lo es, ni en el derecho ni en la economía. Lo mismo que hace que sea alta *PL* —la escala de las operaciones del ferrocarril— hace que *B* también sea alta. La razón de *B* a *PL* no se ve afectada por la escala de la operación del victimario potencial, y tal es la razón que nos permite diferenciar entre los cuasidelitos intencionales y los no intencionales en un sentido económicamente relevante.

Supongamos que quiero un automóvil y decido obtenerlo robando el automóvil de usted. *B* no es sólo menor que en un caso de accidente: es en realidad un número negativo, porque en lugar de ahorrar recursos perjudicando a la víctima (lo que implica una *B* positiva) ahorraría recursos no perjudicando a la víctima (lo que implica una *B* negativa), ya que deberá costarme algo el

terés se protege al convertir un asalto en un cuasidelito, dado que la víctima puede recuperar los daños causados por el asalto aunque no se vea efectivamente atemorizada por el gesto? Existe la interferencia con la paz mental, de forma análoga a la causada por una petición telefónica no deseada, incluso si no hay resultados reales de temor. También puede ser que se castigue el asalto como una forma de prevenir la conducta que es probable que se traduzca en un perjuicio real. Esto sería una excepción al principio de que los actos preparatorios sólo se castigan penalmente, y no por el derecho aplicable a los cuasidelitos. Véase la discusión en el siguiente capítulo y Richard A. Posner, *The Economics of Justice,* 285-286 (1981).

[81] Es improbable que el ladrón le fije al bien un valor mayor que el que le fija el propietario. Si el ladrón lo valúa más alto, es probable que también lo hagan otros, algunos de los cuales tienen los recursos necesarios para comprarlo; y el propietario estará encantado de vendérselo a uno de ellos.

robo del automóvil. Por supuesto, hay una ganancia por el otro lado, porque de otro modo no robaría el automóvil, pero esa ganancia no representa un beneficio social neto, porque la contrarresta la pérdida del automóvil para la víctima. Además, *P* es muy elevada —mucho mayor que en un caso de accidente—, porque el hecho de querer causar un perjuicio a alguien vuelve mucho más probable la ocurrencia de un perjuicio que en el caso en que el perjuicio ocurre simplemente como un subproducto no deseado de otra actividad como la de transportar carga de un punto a otro. Por lo tanto, *B* no es sólo menor que *PL* en un caso de un delito intencional, como lo es en un caso de un delito negligente, sino que es drásticamente menor.

Esto tiene dos implicaciones importantes para la política legal:

1. Esperaríamos, y encontramos, que el derecho está mucho más dispuesto a otorgar daños punitivos en los casos "reales" de cuasidelito intencional que en los casos que carezcan de las características, ya descritas, de tal caso, es decir, que no impliquen una transferencia coercitiva pura. Sabemos que en un caso de responsabilidad estricta los daños punitivos conducirían a una disuasión excesiva. Menos obviamente, lo mismo se aplica en un caso de negligencia simple. En virtud del error judicial y del componente de negligencia que posee la responsabilidad estricta, no puede evitarse por completo la negligencia gastando *B* en cuidado. Por lo tanto, si se aumenta artificialmente a *PL* añadiendo a *L* daños punitivos, los victimarios potenciales se verán inducidos a gastar más que *B* en la prevención de los accidentes, lo que es ineficiente. Pero, dado que la brecha existente entre *B* y *PL* es mucho mayor en el caso "real" de cuasidelito intencional, se minimiza el peligro de disuadir la conducta socialmente valiosa otorgando un pago de daños mayores que *L* y otras políticas aparecen en escena, como la de asegurarnos de que el pago de indemnizaciones sea un disuasivo eficaz al resolver todas las dudas acerca de los daños efectivamente sufridos por el demandante a favor de éste; esto puede hacerse añadiendo un poco de daños punitivos a la estimación de sus daños efectivos.

Además, dado que queremos canalizar la asignación de los recursos a través del mercado en la mayor medida posible, no se me debe permitir sentirme indiferente entre el robo y la compra del automóvil de mi vecino. Por lo tanto, el pago de daños y perjuicios se debe hacer mayor que el valor del automóvil para que yo no considere la apropiación ilícita como un sustituto aceptable de la compra. Los daños y perjuicios punitivos son un procedimiento para lograrlo. Otro procedimiento, también común en los casos del cuasidelito intencional, consiste en hacer que el cuasidelincuente pague a la víctima lo que la cosa robada valía para el cuasidelincuente, antes que lo que valía para la víctima (si es menos). Ésta es la medida restitutiva de los daños mencionada en el capítulo IV. Se utiliza en casos de cuasidelito intencional para tratar de hacer que el cuasidelito carezca de valor para el cuasidelincuente y canalizar así la asignación de recursos a través del mercado.

Pero la medida restitutiva es inadecuada cuando la probabilidad de aprehender al cuasidelincuente es menor a 1. El pago de daños compensatorios también es insuficiente. Si se encubre un cuasidelito (como en un accidente en el que el conductor se da a la fuga), se deben sumar el pago punitivo por daños o la sanción penal al beneficio del demandado o a la pérdida de la víctima para proporcionar la disuasión adecuada. Supongamos que *PL* es de 100, pero la probabilidad de que se capture al cuasidelincuente si comete el cuasidelito y se le obligue a pagar *L* es sólo de 50%; entonces enfrenta un costo esperado del juicio de sólo *PL*/2, por lo que el otorgamiento del pago de daños compensatorio debe duplicarse para traer el costo esperado del juicio de vuelta a *PL* (es decir, $2 \times 0.5PL = PL$).

2. No hay ninguna razón para permitir una defensa de negligencia contribuyente en un caso "real" de cuasidelito intencional (transferencia coercitiva pura), ya que el costo de evitarlo es obviamente menor para el victimario que para la víctima; en efecto, es negativo para el victimario y positivo para la víctima. La víctima *no puede* ser quien pueda evitar el daño a un costo menor.

Hay algunos casos intermedios entre la transferencia coercitiva pura, con su *B* negativa, y el caso ordinario de la negligencia, con su *B* positiva aunque menor a *PL*, o un caso de responsabilidad estricta donde *B* puede ser efectivamente mayor a *PL*. Veamos por ejemplo el caso de la falta de cuidado. Decido descansar la vista mientras voy manejando y me estrello a gran velocidad contra un grupo de peatones. *B* es positiva pero en extremo baja, mientras que *P* y *L* son en extremo elevadas. Aquí opera el argumento de la "resolución de las dudas" para el pago punitivo por daños y perjuicios, al igual que el argumento en contra de que se permita una defensa de la negligencia contribuyente. Por lo tanto, no nos sorprende descubrir que en tal caso permite la ley el otorgamiento de daños punitivos y no permite una defensa de negligencia contribuyente.

Hasta aquí, en nuestra discusión de los cuasidelitos intencionales hemos hecho hincapié en los cuasidelitos adquisitivos: la apropiación ilícita, el robo (apropiación ilícita más asalto en términos de los cuasidelitos), el asesinato por dinero, etc., pero algunos cuasidelitos intencionales tienen una motivación diferente: las utilidades interdependientes. En el capítulo V encontramos las utilidades positivas interdependientes. Aquí nos interesan las utilidades negativas interdependientes. Para humillar a B, que es su enemigo, A le escupe en la cara. No hay ninguna transferencia de riqueza, pero la utilidad de A aumenta al disminuir la utilidad de B. Si el aumento de la utilidad de A es mayor que la disminución de la utilidad de B, la transacción maximiza la utilidad. Pero un tribunal no podría determinar las utilidades relativas; y la desutilidad agregada de permitir este tipo de prácticas —los temores y las medidas de protección que se incitarían— sería muy grande. También hay una cuestión acerca de si se debe dar el mismo peso a la utilidad del cuasidelincuente que a la utilidad de su víctima (véase VII.1 después).

El ejemplo del escupitajo sugiere otra razón económica para otorgar pago punitivo por daños en algunos casos del cuasidelito: aliviar la presión sobre el sistema de justicia directa e indirectamente proveyendo un sustituto de la autoayuda violenta (criminal en sí misma). La expectativa de los daños y perjuicios punitivos provee un incentivo para que las víctimas de una conducta criminal menor soporten el costo de la acción legal contra los perpetradores, sustituyendo el castigo penal por un remedio de cuasidelito. Sin un remedio de cuasidelito significativo (que requiera indemnización punitiva si los daños compensatorios fuesen pequeños o imposibles de probar o si el demandado es rico y no se le disuadiría de conducirse de ese modo mediante la posibilidad de tener que pagar daños modestos), es más probable que las víctimas de la conducta de tipo criminal demasiado menor para despertar el interés del fiscal de distrito se venguen directamente de su asaltante, y un propósito histórico de la ley es impedir las alteraciones del orden. La Corte Suprema ha dicho que "pocos otorgamientos [de daños punitivos] superiores a una razón de un dígito entre el pago punitivo por daños y el pago de daños compensatorio, en un grado significativo, satisfarán el debido proceso".[82] El Tribunal continuó sugiriendo que "cuatro veces el monto del pago de daños compensatorio podría estar cerca de la línea de impropiedad constitucional". Existen buenas razones para desear poner algunos límites a la concesión del pago punitivo de daños. El castigo excesivo puede disuadir en exceso; el demandado debe tener una notificación razonable de la sanción por actos ilegales para que pueda tomar una determinación racional de cómo actuar; y las sanciones se deben basar en el mal hecho mas no en la condición del demandado: se debe castigar a la persona por lo que hace, no por quien es, incluso si el tal es una gran corporación. Ésta es la esencia de la teoría aristotélica de la justicia correctiva, a la que, como veremos en el capítulo VIII, se puede dar una interpretación económica.

Sin embargo, lo que sigue es que debe asignarse el pago punitivo por daños a través de normas antes que de una manera completamente *ad hoc*, y esto no nos dice cuál debe ser la razón máxima del pago punitivo al pago compensatorio por daños en un caso particular. El límite que sugiere la Corte Suprema tiene sentido cuando los daños compensatorios son muy grandes; porque, si son pequeños, limitar a un múltiplo de un solo dígito el pago punitivo por daños podría no lograr los fines para otorgar el pago punitivo por daños.[83] Supongamos que los daños compensatorios en el caso del escupitajo fueron de 50 dólares, de modo que si el pago punitivo por daños no podrá exceder un múltiplo de un solo dígito de daños compensatorios el otorgamiento del pago máximo por daños sería de 500 ($50 + (9 \times 50)$). Ésa sería una suma demasiado pequeña para motivar la presentación de lo que podría ser una demanda muy meritoria. Tendría más sentido económico que el

[82] State Farm Mutual Automobile Ins. Co. *vs*. Campbell, 123 S. Ct., 1513, 1524 (2003).
[83] Mathias *vs*. Accor Economy Lodging, Inc., 347 F.3d, 672 (7th Cir., 2003).

monto del pago punitivo por daños sea *inverso* a la cantidad de daños compensatorios.[84] Porque, entre menores sean los daños compensatorios, las víctimas tendrán menos incentivos para demandar y, por tanto, será mayor la necesidad de que los daños punitivos creen un incentivo para demandar, y así impidan la conducta ilícita.

La riqueza del demandado debe figurar en el cálculo del pago punitivo por daños si su riqueza le permite a él (es más probable que sea a ella: una corporación u otra institución) preparar una defensa ferozmente agresiva contra las demandas que tienen inversiones pequeñas y al hacer esto volver muy costoso el litigio de estos casos en su contra. Esto, a su vez, puede hacer que para un demandante sea difícil encontrar a un abogado dispuesto a llevar su caso sobre la base habitual de la tarifa contingente de 33-40%. En otras palabras, el demandado está invirtiendo en el desarrollo de una reputación que pretende disuadir a los demandantes. Un otorgamiento estricto del pago punitivo por daños mermaría esa estrategia.

Una cuestión sutil de intencionalidad deriva del concepto de la "indiferencia deliberada", que desempeña un papel especialmente importante en muchos cuasidelitos constitucionales tales como la imposición de castigos crueles e inusitados. Los prisioneros se quejan con frecuencia de que las condiciones de salud y seguridad de la prisión son tan desastrosas que debiera considerarse que los funcionarios de la prisión están imponiendo un castigo. Los tribunales insisten en que la negligencia de los funcionarios de la prisión sea intencional, pero permiten que la "indiferencia deliberada" hacia la salud y la seguridad de los prisioneros satisfaga este requerimiento. Este término significa que los demandados están conscientes de las condiciones peligrosas o nocivas y deciden no hacer nada al respecto. Como señalamos al principio de esta sección, las empresas grandes que realizan actividades peligrosas tales como la construcción de puentes o túneles "saben" que sus actividades implicarán un accidente ocasional y "deciden" no hacer nada al respecto (lo que requeriría el abandono de sus actividades) sin ser declarados cuasidelincuentes intencionales. Pero esto sólo significa que la "indiferencia deliberada" es una condición necesaria antes que, como podría suponerse que implica la fórmula judicial, una condición suficiente de la responsabilidad como un cuasidelincuente intencional. Entre mayor sea *PL* (especialmente *P*), más probable será que el victimario potencial sepa que su conducta es peligrosa; en consecuencia, el conocimiento se convierte en una aproximación a una *PL* elevada, y, en igualdad de las demás circunstancias, entre más elevada sea *PL* mayor será la falta del demandado. Además, el componente de los costos de información de *B* es menor si el demandado estaba efectivamente consciente del riesgo y decidió no hacer nada al respecto. Por lo tanto, la indiferencia deliberada aumenta la probabilidad de una gran brecha entre *PL* y *B*.

[84] Véase Kemezy *vs*. Peters, 79 F.3d, 33 (7th Cir., 1996).

CUADRO VI.3. *Clases de casos de cuasidelitos*

Caso	*P*	*L($)*	*PL($)*	*B($)*
Inevitables	0.002	$10 000	$20	$20 000
No negligentes	0.002	10 000	9 000	10 000
Negligentes	0.002	10 000	20	15
Extremadamente negligentes	0.002	10 000	20	5
Imprudentes	0.800	10 000	8 000	2
Intencionados	0.900	10 000	9 000	100

El cuadro VI.3 ofrece valores numéricos hipotéticos de *P, L* y *B* para ilustrar los grados de la culpabilidad. En las primeras cuatro hileras, $PL < B$, por lo que no existe responsabilidad; en las hileras restantes, $PL > B$, por lo que sí existe responsabilidad. La última hilera, titulada "intencionados", es para el caso en el que *B* se vuelve positivo porque el cuasidelincuente, desistiendo del cuasidelito, perdería la satisfacción que le habría de conferir el sufrimiento de su víctima (de modo que la seguridad sería costosa para él); éste es el caso de las utilidades negativas interdependientes.

La indiferencia deliberada no se manifiesta por separado. En la práctica, no se distingue de la imprudencia, ya que, si el riesgo es muy grande en relación con los costos de su evitación, es muy probable que un juez o un jurado infieran que el demandado estaba consciente del riesgo e, inexcusablemente, no hizo nada al respecto.

VI.16. LA DIFAMACIÓN

Clasificada de ordinario como un cuasidelito intencional, porque el hecho de escribir o hablar críticamente acerca de una persona es un acto deliberado, la difamación tiene un fuerte sabor de responsabilidad estricta, como ocurre con la regla de que no es una defensa el hecho de que el demandado haya hecho un esfuerzo razonable para no difamar al demandante. En Jones *vs.* E. Hulton & Co.,[85] el autor de un relato periodístico ficticio le puso accidentalmente a uno de los personajes el nombre de una persona real, Artemus Jones, quien demandó por difamación y ganó al demostrar que sus vecinos pensaban que el relato se refería a él. Jones no podía evitar el perjuicio, y, aun si el periódico hubiese sido cuidadoso, la literatura de ficción disfrazada de verdad no es una actividad tan valiosa que no se hubiese podido evitar el perjuicio a un costo razonable mediante un cambio de esa actividad. Ade-

[85] [1909] 2 K. B., 444, *aff'd* [1910] A. C., 20.

más, el incidente tenía un sabor de temeridad. Como un hombre que dispara a las ventanas de una casa que cree deshabitada y accidentalmente mata a alguien, el periódico estaba difamando a quien creía que era una persona inexistente, pero alguien con el nombre y algunas características del personaje ficticio sería perjudicado por el ataque, y lo fue. Con todo, el *distribuidor* del periódico no habría sido responsable ante Jones a menos que fuese negligente al no detectar el libelo; porque resulta difícil precisar cuál cambio en el nivel de la actividad sería óptimo para el distribuidor.

El derecho de la difamación tiene dos fallas desconcertantes que la economía podría ser capaz de explicar. La primera es que los libelos grupales (por ejemplo, "Todos los médicos son charlatanes") no generan una acción. No es probable que el libelo perjudique a los miembros del grupo porque la sustitución de toda una ocupación es mucho más costosa que la sustitución de un individuo. Si el libelo fuese "El doctor Jones es un charlatán", los pacientes de Jones podrán cambiar fácilmente a otros médicos. ¿Pero qué pueden hacer si todos los médicos son charlatanes? Igual podrían quedarse con el viejo Jones, a quien por lo menos conocen. La segunda regla, extrañamente relacionada con la anterior, es que no puedes difamar a los muertos (con una excepción que se advierte más adelante). La difamación es un perjuicio contra la reputación, y como vimos al discutir el cuasidelito de la invasión de la privacidad en el capítulo III, la reputación es una base para inducir a otros a realizar transacciones con uno dentro o fuera del mercado. Si estás muerto y, por lo tanto, has dejado de hacer transacciones, el perjuicio ha cesado. Pero si el libelo fuera que el occiso había tenido una enfermedad heredable, las consecuencias adversas no terminarían con la muerte, y el derecho reconoce una excepción en tales casos.

El derecho trata a las difamaciones por escrito (libelos) con mayor severidad que a las difamaciones orales (calumnias). El costo de la prevención de la difamación es menor en una comunicación escrita que en una oral. Escribir es una actividad más deliberativa; hay tiempo para considerar la posible influencia de nuestras palabras, algo que a menudo no existe cuando se habla. Por lo tanto, es menor el peligro de que la responsabilidad por la difamación impida las comunicaciones socialmente valiosas que se hacen por escrito que en el caso de las que se hacen oralmente. Esto implica —y hay un apoyo judicial para tal posición— que un discurso por radio o televisión que se lee de un guion, en vez de pronunciarlo de modo improvisado, deberá clasificarse como libelo y no como calumnia. Igual sucede si se improvisa, pero a partir de notas. El orador habrá tenido tiempo para ordenar sus pensamientos y, por lo tanto, debiera haber sido capaz de evitar el cometer inadvertidamente la difamación.

Consideremos ahora la defensa del privilegio, por ejemplo, el privilegio que tiene el empleador para dar una carta de referencia a un empleado que podría contener material difamatorio. El beneficio de una referencia franca

afectaría al empleador potencial antes que al actual, y sería difícil que el primero compensara al segundo. Aunque en principio el primero podría prometer que indemnizaría al segundo por cualesquiera daños y costos resultantes de un juicio por difamación, los beneficios son demasiado triviales para que tales arreglos valgan la pena. La respuesta del derecho consiste en externalizar los costos de una carta de referencia a fin de estimular el otorgamiento de un beneficio externo al empleador potencial. El privilegio no es absoluto; desaparece si el empleador que otorga la carta de referencia sabía que tal carta contenía acusaciones falsas contra el empleado. En tal caso, no se confiere ningún beneficio externo, de modo que tampoco hay razón alguna para permitir que el empleador externalice un costo. Pero la verdad es una defensa absoluta porque hay un beneficio externo del discurso genuino, aunque el demandado pensara que estaba mintiendo.

Si un periódico publica un artículo difamatorio, y un segundo periódico sólo reporta el contenido de ese artículo (quizá con un prefacio de "El *Times* ha reportado que…"), ¿debiera ser responsable el segundo periódico o sólo el primero? Ambos debieran ser y son responsables. De otro modo cada periódico se sentiría reacio a ser el primero en publicar un artículo controversial que podría originar una reclamación por difamación. Cada periódico desearía ser un segundo cercano a fin de evitar el costo de ser demandado, sin que le ganaran malamente la delantera. Si todos los editores de periódicos pensaran de este modo, podría haber una demora considerable en la publicación de artículos controversiales. Si el editor original y los reproductores son responsables por cualquier difamación contenida en el artículo original, ya no hay ningún incentivo para que un editor se quede atrás; en consecuencia, las noticias se hacen públicas más rápido. Incluso si no existe una contribución entre los cuasidelincuentes conjuntos, el fallo esperado contra el editor original será menos severo y, por lo tanto, el incentivo para publicar primero será mayor que bajo un sistema en el que los editores subsecuentes sean inmunes contra toda demanda.

La difamación de una figura pública, como un político o un atleta profesional famoso, presenta una cuestión de libertad de expresión que se discute después en XXIX.2.

VI.17. La responsabilidad del sucesor

El aumento del conocimiento científico ha reforzado la conciencia de las consecuencias adversas que, con una sustancial demora, tienen actividades tales como la venta de asbestos o DES (dietilestilbestrol). Entre mayor sea el intervalo entre una actividad y sus consecuencias, mayor será la probabilidad de que el actor ya no exista cuando se presenten las demandas por daños. Entonces la cuestión se trata de si los sucesores del actor debieran ser

responsables. Otro planteamiento de esta cuestión, que identificará a una de las consideraciones económicas críticas, consiste en preguntar si los activos debieran cargar una responsabilidad contingente. Supongamos que una corporación fabricante de asbestos tenía muchos años de haber vendido sus activos, distribuido el dinero entre los accionistas y luego se había disuelto. ¿Debieran los sucesores ser responsables ante las víctimas de la asbestosis (quizás hasta la cantidad distribuida entre ellos al disolverse la corporación, a fin de preservar el principio de la responsabilidad limitada que veremos en el capítulo XV)? ¿O los compradores de los activos? Si nadie es responsable, la corporación habrá logrado externalizar los costos sanitarios de los asbestos. Pero si los accionistas o los compradores son responsables, los costos de las transacciones corporativas serán mayores.

Las pérdidas sufridas y ganancias cosechadas hacen de tales medidas un dilema difícil, pero pueden identificarse un par de factores relevantes. Uno es si se previó la responsabilidad cuasidelictiva de la corporación en el momento de la disolución. Si se previó, la responsabilidad del sucesor no será tan costosa como habría sido si no se la hubiese previsto, porque debiera ser posible estimar su costo y ajustar consiguientemente el precio de compra. La responsabilidad del sucesor será también menos costosa si la corporación se vende en una sola entrega en lugar de que sus activos se dispersen entre muchos compradores, porque entonces el número de juicios de sucesión se reducirá de muchos a uno solo.

La regla general del derecho común es que las entidades sucesoras no son responsables de los cuasidelitos de sus predecesores, a menos que la responsabilidad se asuma expresamente en el contrato de venta. Esa regla refleja una época en la que las consecuencias demoradas del cuasidelito eran menos comunes que en la actualidad, de modo que los costos de transacción de la responsabilidad de los sucesores dominaban la externalización de los costos de los accidentes. Siendo ahora más comunes las consecuencias demoradas, la regla se está erosionando, como el análisis económico pronosticó que lo haría.

Una venta de activos no es el único procedimiento para evitar la responsabilidad plena de los cuasidelitos; otro consiste en asegurarse de que nunca habrá activos suficientes para pagar un fallo enorme por concepto del cuasidelito. Ésa parece ser una receta para la quiebra, pero considérese lo siguiente: supongamos que el costo de eludir un accidente de 40 millones serían 200 000 dólares, y que la probabilidad del accidente es de 0.01 si no se gasta nada tratando de evitarlo, de modo que todo gasto hasta por 400 000 dólares que impidiera el accidente estaría justificado por su costo desde un punto de vista social. Supongamos también que, a un costo relativamente bajo (y supongamos que tal costo es cero), la empresa puede evitar tener jamás más de un millón en activos que pudieran embargarse para pagar un fallo legal. Entonces, el costo esperado del fallo para la empresa es de sólo 10 000 dólares

(0.01 × un millón), y la empresa no gastará 200 000 dólares en la prevención de accidentes. Congruente con este análisis, se ha observado que los proyectos industriales riesgosos se concentran desproporcionadamente en empresas pequeñas, porque su responsabilidad potencial por cuasidelitos se trunca como en el ejemplo.[86]

VI.18. Exenciones[87]

Supongamos que se acusa a un paramédico de negligencia al no poder salvar la vida de un paciente.[88] Incluso si se le declarara negligente no sería responsable, porque se exenta al personal médico de emergencia de la responsabilidad por negligencia simple; su conducta se debe declarar "voluntaria y gratuita" (temeraria) para crear la responsabilidad por el cuasidelito. Es probable que la responsabilidad por negligencia simple disuada aun a los rescatistas competentes de ofrecer sus servicios, ya que si el rescate falla podrían enfrentar una demanda. El problema es especialmente grave si, como generalmente será el caso, el rescatista no puede buscar la compensación por el beneficio otorgado por un rescate exitoso. No queremos que el paramédico compare el costo probable para él de un rescate que falle y se traduzca en una demanda en su contra con los beneficios para la persona que necesita ser rescatada: queremos que siga adelante y trate de rescatar a la persona. Sólo se impone la responsabilidad si responde temerariamente, una respuesta con mayor probabilidad de perjudicar a la persona que de ayudarla, quizá por interponerse en el camino de un rescatista competente.

Por razones similares, a la policía y a otros funcionarios o representantes del gobierno se les otorga con frecuencia lo que se denomina inmunidad oficial, que los protege de la responsabilidad de los cuasidelitos por los perjuicios derivados de actos oficiales (como la detención de los presuntos delincuentes), siempre que, de nuevo, no actúen imprudentemente. Sin tal inmunidad podrían tener escrúpulos acerca del cumplimiento de sus funciones, porque estarían poniendo en una balanza los riesgos legales para sí mismos, de un lado, y los beneficios sociales del cumplimiento de sus funciones, del otro.[89]

[86] Al H. Ringleb y Steven N. Wiggins, "Liability and Large Scale, Long Term Hazards", 98 *Journal of Political Economy*, 574, 1990.

[87] Gerrit de Geest, "Who Should Be Immune from Tort Liability?", 41 *J. Legal Stud.*, 291 (2012).

[88] Como en Fagocki *vs*. Algonquin Fire Protection Dist., 496 F.3d, 623 (7th Cir., 2007).

[89] Véase más adelante sobre la inmunidad caritativa (actualmente abolida, en gran medida) a la responsabilidad por los cuasidelitos en VIII.3.

Lecturas recomendadas

Arlen, Jennifer H., "Compensation Systems and Efficient Deterrence", 52 *Md. L. Rev.,* 1093 (1993).

_____, "Contracting over Liability: Medical Malpractice and the Cost of Choice", 158 *U. Pa. L. Rev.*, 957, 987 (2010).

Calabresi, Guido, *The Costs of Accidents: A Legal and Economic Analysis* (1970).

Croley, Steven P., y Jon. D. Hanson, "Rescuing the Revolution: The Revived Case for Enterprise Liability", 91 *Mich. L. Rev*., 683 (1993).

Danzon, Patricia M., *Medical Malpractice: Theory, Evidence and Public Policy* (1985).

Friedman, David, "An Economic Explanation of Punitive Damages", 40 *Ala. L. Rev*., 1125 (1989).

King, Elizabeth M., y James P. Smith, "Computing Economic Loss in Cases of Wrongful Death" (RAND Inst. for Civil Justice, R-3549-ICJ, 1988).

Landes, William M., y Richard A. Posner, *The Economic Structure of Tort Law* (1987).

Polinsky, A. Mitchell, y Steven Shavell, "Punitive Damages: An Economic Analysis", 111 *Harv. L. Rev*., 869 (1998).

_____, "The Uneasy Case for Product Liability", 123 *Harv. L. Rev.*, 1437 (2010).

Porat, Ariel, y Alex Stein, *Tort Liability under Uncertainty* (2001).

Quah, Euston, y William Rieber, "Value of Children in Compensation for Wrongful Death", 9 *Int'l L. Rev. Law and Econ.*, 165 (1989).

Rubin, Paul H., y Joanna M. Shepherd, "Tort Reform and Accidental Deaths", 50 *J. Law & Econ.*, 221 (2007).

Shavell, Steven, *Economic Analysis of Accident Law* (1987).

Sykes, Alan O., "The Boundaries of Vicarious Liability: An Economic Analysis of the Scope of Employment Rule and Related Legal Doctrines", 101 *Harv. L. Rev*., 563 (1988).

"Symposium, Current Research on Medical Malpractice Liability", 36 *J. Legal Stud.* S1 (2007).

The Cost of Libel: Economic and Policy Implications (Everette E. Dennis y Eli M. Noam [comps.], 1989).

Tort Law and Economics (Michael Faure [comp.], 2009).

Problemas

1. Evalúe la proposición siguiente: La medida estándar de los daños en los casos de cuasidelito es equivalente al estándar contractual de los daños de

confianza y, por lo tanto, se traduce en una desistencia insuficiente de la conducta cuasidelictiva.

2. Evalúe la enunciación siguiente desde el punto de vista de la economía:

> En efecto, una de las mayores preocupaciones prácticas generadas por el aumento de las primas de seguros contra la mala práctica es la incapacidad de los médicos jóvenes que inician su práctica para pagar sus costos de seguros, que son, después de todo, sólo uno de los muchos costos que deben cubrirse con el ingreso bruto de su práctica.
>
> En general, se puede afirmar con seguridad que se ha exagerado la capacidad de los médicos y las compañías de seguros para revertir los costos de las cada vez mayores pérdidas por la mala práctica. En la medida en que estas pérdidas surgen de incidentes de años anteriores, no hay manera de que los médicos o las compañías de seguros recuperen estos "costos hundidos" cobrando honorarios elevados por sus servicios futuros […]
>
> Richard A. Epstein, "Medical Malpractice: The Case for Contract", 1 *Am. Bar Foundation Res. J.*, 87, 88 n. 3 (1976).

3. Melvin M. Belli, "The Use of Demonstrative Evidence in Achieving 'the More Adequate Award'", Mississippi State Bar Association, 33 (1951), cuestiona la práctica del descuento de los ingresos futuros hasta el valor presente. Belli argumenta que los ingresos del demandante debieran multiplicarse por el periodo de su incapacidad y luego multiplicarse de nuevo para reflejar el aumento estimado en el costo de la vida durante el periodo de la incapacidad. ¿Es éste el enfoque económico correcto? ¿Cómo argumentaría en contrario ante un juez o un jurado?

4. Un avión se estrella a resultas de un defecto no negligente en un instrumento provisto por A. El avión fue fabricado por B con diversos componentes, incluido el instrumento provisto por A, y es operado por C, una aerolínea. ¿Cuál diferencia económica hay, de haberla, si A es responsable de los daños resultantes del accidente, o si es responsable B, o C? Véase Goldberg *vs.* Kollsman Instrument Corp., 191 N.E. 2d 81 (N.Y., 1963); IV.5 antes.

5. Supongamos que hay una clase de accidentes automovilísticos en la que todos los victimarios son hombres muy ricos. Los costos de oportunidad de su tiempo son tan grandes que los costos esperados de los accidentes son menores a los costos para ellos de prevenir los accidentes conduciendo más lentamente, siendo la conducción más lenta el único método por el que podrían prevenirse los accidentes. En estas circunstancias, ¿cuál regla de responsabilidad sería más eficiente: la responsabilidad estricta, sin defensa de la negligencia contribuyente, o la no responsabilidad en absoluto?

6. ¿Puede argumentarse que hay una justificación económica mayor para la responsabilidad estricta por los productos para los transeúntes que tal responsabilidad para los compradores del producto?

7. Supongamos que un médico se topa con una persona lesionada que yace en la calle sufriendo gran dolor. El médico insiste, como una condición para atenderlo, en que la persona lesionada renuncie a la responsabilidad del médico por cualquier reclamación de mala práctica derivada del tratamiento. ¿Debieran los tribunales respetar la renuncia?

8. Si el sistema de los cuasidelitos se considerara como un método de la seguridad social antes que un método de disuasión de los accidentes no económicos, ¿serían las condenas en los casos de cuasidelitos mayores o menores en promedio que bajo la justificación de la desistencia?

9. Se ha observado a menudo que cuando los mineros se ven atrapados en una mina de carbón, o cuando un alpinista se pierde en la cima de una montaña, la sociedad destinará recursos enormes a la salvación de sus vidas, recursos aparentemente desproporcionados a los gastos de la sociedad para prevenir que los individuos se pongan en peligro por su cuenta. Explique la paradoja.

10. Este capítulo señala que la víctima potencial de un accidente tiene por lo común derecho a suponer que los victimarios potenciales están utilizando el cuidado debido, de modo que no necesita tomar precauciones que serían óptimas sólo si los victimarios potenciales se estuvieran comportando negligentemente. Pero hay una excepción para el caso en que el peligro es palpable, como ocurre cuando un viajero en un cruce, aunque ve que un tren se aproxima a toda velocidad, no toma más precauciones que las que tomaría para impedir la lesión causada por un tren operado con seguridad. Explique la economía de esta excepción y compárela con la doctrina de la última oportunidad clara.

11. Discuta los pros y contras económicos del hecho de hacer a los anfitriones sociales responsables de los accidentes causados por los invitados que han bebido demasiado en la fiesta del anfitrión. Véase "Note, Turn Out the Lights; the Party's Over: An Economic Analysis of Social Host Liability", 6 N. *Ill. U. L. R. Rev.,* 129 (1986).

12. Aparte de la consideración mencionada antes, en VI.12, ¿debieran estar sujetos a impuesto los otorgamientos de daños en los casos de lesiones personales? *Cf.* Caterpillar, Inc. *vs.* Lewis, 519 U.S., 61 (1996).

13. Vimos que hay tres métodos para evaluar los servicios de un ama de casa: calcular el costo de reposición, fijar el costo de oportunidad y determinar el monto de la producción total de la unidad familiar que es aportado por la esposa. (Los tres métodos se discuten, desde el punto de vista de la economía forense —es decir, el uso de la economía por los peritos en un juicio—, en Charles C. Fischer, "The Valuation of Household Production: Divorce, Wrongful Injury and Death Litigation", 53 *Am. J. Econ. & Sociology,* 187 (1994)). ¿Cómo esperaría que difirieran en la práctica?

14. Supongamos que, como una alternativa al otorgamiento del pago de daños para la víctima de un cuasidelito, la ley le otorgara la ganancia que el

victimario hubiera obtenido por lesionarlo (el método restitutivo de medición del alivio monetario). ¿Cuál sería la justificación económica de tal enfoque?, y ¿sería una alternativa eficiente al pago de daños y perjuicios convencional?

15. ¿De qué manera, si acaso, debieran cambiar las doctrinas del derecho aplicable a los cuasidelitos a la luz de la comprensión actual (a la que preceden las doctrinas) de la psicología humana relevante? Véase Michael G. Faure, *The Impact of Behavioural Law and Economics on Accident Law* (2009); Barbara Luppi y Francesco Parisi, "Behavioral Models in Tort Law", *Social Science Research Network*, 2012, http://papers.ssrn.com/so13/papers.cfm?abstract_id=2165856.

VII. EL DERECHO PENAL[1]

VII.1. La naturaleza y la función económicas del derecho penal

Los tipos del comportamiento ilícito examinados en los capítulos anteriores, principalmente los cuasidelitos y los incumplimientos de contrato, sujetan al actor que delinque al pago de daños monetarios a su víctima, o a veces a la prohibición (es decir, impuesta por fallo), so pena de ser sancionado por desacato al tribunal, de continuar o repetir el ilícito, pero en ambos casos sólo si la víctima demanda. No obstante, los delitos son perseguidos por el Estado, y el delincuente es obligado a pagar una multa a éste o a sufrir una pena no pecuniaria tal como la de prisión. El procedimiento judicial es también diferente en las dos clases de casos, pero el examen de las diferencias procedimentales se difiere hasta el capítulo XXII. Por ahora nos interesa descubrir por qué debiera haber sanciones distintas, buscadas por el Estado, para algunos tipos de ilícitos y cuáles doctrinas sustantivas implican tales sanciones.

Hay cinco tipos principales de la conducta ilícita que se consideran delictivos en el régimen jurídico de los Estados Unidos.

1. Los cuasidelitos intencionales, examinados en el capítulo anterior, que representan una transferencia coercitiva pura ya sea de riqueza o de utilidad de la víctima al victimario. El homicidio, el robo simple, el robo a casa o con allanamiento, el latrocinio, la violación, el asalto y la agresión, la mutilación, el robo por falsas pretensiones y la mayoría de los demás delitos que, como éstos, eran punibles de acuerdo con el derecho común inglés corresponden a cuasidelitos intencionales tales como el asalto, la agresión, el allanamiento y la apropiación ilícita, aunque veremos más adelante que los requerimientos del estado mental del demandado y la presencia de un perjuicio real difieren a veces de la contrapartida delictiva del cuasidelito intencional. Sin embargo, veamos ahora algunos ejemplos más problemáticos del delito como transferencia coercitiva pura.

(1) *La falsificación* puede verse como una forma del robo por engaño, siendo éste que el pagador está pagando con moneda legal. Si se descubre la falsificación, la víctima es quien acaba teniendo la moneda sin valor. Si no se

[1] Por lo que toca a las reglas y los principios, respectivamente, del derecho penal, véanse Wayne R. LaFave, *Criminal Law* (4a ed., 2003), y Ronald N. Boyce, Donald A. Drips y Rollin M. Perkins, *Criminal Law and Procedure: Cases and Materials* (9a ed., 2004) (ambos sobre las reglas); y H. L. A. Hart, *Punishment and Responsibility: Essays in the Philosophy of Law* (1968), y George P. Fletcher, *Rethinking Criminal Law* (2ª ed., 2000) (ambos sobre los principios).

descubre, la pérdida se difunde más ampliamente. Dado que la cantidad de dinero en circulación es ahora mayor que antes de la falsificación en relación con el acervo total de bienes y servicios en la sociedad, el dinero de todos vale menos (inflación); todos pierden, menos el falsificador y el deudor cuya obligación de pago está en dólares fijos. Más importante aún —porque, a menos que la falsificación se generalice, el efecto sobre la oferta de dinero será menor—, la falsificación impone los costos de peso muerto en la forma de las herramientas y el tiempo del falsificador (y de la pérdida resultante de empleo productivo) y de las medidas de autoprotección por parte de comerciantes, bancos y otras empresas que reciben dinero del público.

(2) *La violación* elude el mercado de las relaciones sexuales (maritales y de otra índole) del mismo modo que el robo elude los mercados de los bienes y servicios ordinarios. Pero algunos violadores obtienen un placer extra del hecho de que la mujer no haya consentido. Para estos violadores no hay un sustituto en el mercado —los costos de transacción en el mercado son prohibitivos— y, por lo tanto, un economista podría argumentar que para ellos la violación no es una transferencia coercitiva pura y no debiera castigarse si el placer para el violador (medidas por lo que él estaría dispuesto a pagar —aunque no a la víctima— por el derecho de violar) supera al dolor físico y emocional de la víctima. Hay varias objeciones prácticas, como la de que estos violadores resultan difíciles de distinguir empíricamente de quienes son meros ladrones del sexo; la de que si se les diera rienda libre se induciría a las mujeres a invertir mucho en la autoprotección, lo que a su vez incitaría un gran gasto de los violadores para superar esa inversión, y la de que no existe un método objetivo para medir la utilidad del violador y la desutilidad de la víctima con el fin de decidir cuál es mayor. Pero el hecho de que cualquier tipo de licencia para la violación sea incluso concebible en el marco de la teoría de la maximización de la riqueza que guía al análisis de este libro es una limitación de la utilidad de esa teoría. Lo que genera la posibilidad de una licencia para violar es el hecho de que la utilidad del violador se pondera igual que la utilidad de su víctima. Si se otorgara a la utilidad del violador una ponderación de cero en el cálculo de costos y beneficios, una licencia para violar no podría ser eficiente. Pero la única base para tal ponderación sería un principio ético antes que económico.

(3) *La violación marital*:[2] Hasta hace poco tiempo, el matrimonio era una defensa completa contra una acusación de violación. En una sociedad que premia la virginidad premarital y la castidad marital, el daño fundamental de la violación es la destrucción de esos bienes y no se inflige por la violación marital. No debiera sorprendernos que en estas sociedades sea la seducción de una mujer casada un delito más grave que la violación, pues es más probable que produzca hijos… los que no serán del esposo. En tales sociedades,

[2] Véase Richard A. Posner, *Sex and Reason*, cap. 14 (1992).

además, los servicios principales que aporta una esposa al matrimonio son sexuales y procreativos, y si ella trata de privar a su esposo de estos servicios estará atacando al corazón mismo del matrimonio. Un derecho a exigir algo no implica un derecho a tomarlo por la fuerza, pero puede diluir la impropiedad sentida de la fuerza. Asimismo, entre menor sea la tasa de divorcios habrá menos separaciones; y dar prueba de la falta de consentimiento es más difícil si los miembros de la pareja casada siguen juntos. La excepción a la correlación negativa entre el divorcio y la separación ocurre cuando, como en los países católicos hasta hace poco tiempo, el divorcio estaba prohibido, pero se sustituía por las separaciones formales, a menudo permanentes.

Las razones para no considerar a la violación marital como un delito han perdido fuerza con el paso del tiempo. Por ejemplo, la creciente independencia económica de las mujeres, basada en su acceso mejorado al mercado de trabajo, ha aumentado su demanda de control sobre su capacidad reproductiva, por lo que coordinan el momento de llegada y el número de sus hijos con su carrera en el mercado. Cualquier forma de relación sexual involuntaria afecta a ese control, al menos en principio (porque, desde luego, existen anticonceptivos eficaces), mediante la creación de la amenaza de que un nacimiento interrumpirá la trayectoria laboral de la mujer.

Regresemos ahora a nuestra tipología de la conducta ilegal que la ley vuelve penal.

2. La comisión deliberada de actos que reducen el bienestar social, tales como la fijación de precios (véanse a este respecto los capítulos IX y X), cuya ilicitud podría no haberse reconocido en el derecho común.

3. Los intercambios voluntarios y, por ende, presuntamente (pero sólo presuntamente, como vimos en los capítulos III al VI) maximizadores del valor que involucran actividades que el Estado ha prohibido, tales como la prostitución, la venta de pornografía infantil, la venta de bebés para adopción, la venta de servicios de transportación regulados a precios que no aparecen en las tarifas publicadas por el transportista y el narcotráfico.

4. Ciertos actos preparatorios amenazantes pero que no son cuasidelitos, como el intento o la conspiración infructuosos para asesinar a alguien cuando no están presentes los elementos de una tentativa cuasidelictiva (como no lo estarían, por ejemplo, si la víctima no fue perjudicada y ni siquiera supo de la tentativa en el momento en que se realizó).

5. La conducta que, si se permitiera, complicaría otras formas de la regulación del derecho común. Tenemos ejemplos de esto en el abandono de la escena de un accidente y el ocultamiento fraudulento de activos al acreedor de un fallo judicial.

¿Pero por qué no pueden incluirse las cinco categorías en el derecho aplicable a los cuasidelitos? Surge de inmediato una respuesta para las cate-

gorías *3* y *4*: nadie resulta lastimado. Pero ésta es una respuesta superficial; podríamos permitir que demandara daños punitivos cualquier persona a quien la ley tratara de proteger. Una respuesta mejor, aunque incompleta, es que la detección resulta difícil cuando no hay ninguna víctima que reporte el ilícito y declare contra el delincuente. Los daños punitivos pueden ajustarse hacia arriba para tomar en cuenta la dificultad de la detección; en principio, este instrumento podría aplicarse también a los delitos de la categoría 5. Sin embargo, entre mayor sea el nivel de los daños punitivos, menos probable será que constituyan una sanción viable. Otro interrogante acerca de las categorías 3 y 4 es éste: ¿para qué castigar los actos que no lesionan a nadie? Para la categoría 3, la respuesta se encuentra afuera de la ciencia económica; resulta difícil para un economista entender por qué, si un delito carece verdaderamente de "víctimas", debiera castigarse al criminal. (Por supuesto, los delitos que carecen ostensiblemente de víctimas podrían tener efectos para terceros, como ocurre con otros intercambios contractuales; la venta de licor a un conductor ebrio es un ejemplo.) Para la categoría 4, la respuesta se liga al interrogante —del que podremos ocuparnos ahora— de la causa por la que el derecho aplicable a los cuasidelitos no es adecuado para tratar las categorías 1 y 2 (las transferencias coercitivas que violan los principios del derecho común o de las leyes reguladoras).

En el capítulo anterior vimos que la sanción apropiada para una transferencia puramente coercitiva como el robo es algo mayor que la estimación hecha por la ley de la pérdida de la víctima, donde la diferencia trata de confinar las transferencias al mercado siempre que los costos de transacción del mercado no sean prohibitivos.[3] Así pues, los daños iguales al valor de mercado serían inadecuados, aunque el propietario valuara el bien en cuestión a su precio de mercado. Bien podría valuarlo a un precio mayor (recuérdese que el precio de mercado es el valor para el comprador marginal, no para el comprador medio). La ley no puede medir con facilidad los valores subjetivos, de modo que éste es un argumento a favor de la adición de un bono muy cuantioso a los daños basados en el valor de mercado, es decir, un argumento a favor de fuertes daños punitivos en el caso de un robo.

En el caso de los delitos que causan muertes o aun apenas un riesgo sustancial de muerte, los daños óptimos serán a menudo astronómicos. Como se ve en la gráfica VI.3, la relación entre el riesgo y la compensación plena por asumirlo no es lineal. El riesgo de ser muerto por accidente se distribuye más o menos al azar por toda la población, pero el riesgo de ser asesinado se concentra en el número relativamente pequeño de personas que constituyen obs-

[3] Sujeto a dos calificaciones, que discutiremos más adelante con mayor detalle. La primera es que los delincuentes respondan a las sanciones monetarias mayores. En caso contrario, se complica el diseño de una estructura de penas. En segundo lugar, si son muy elevados los costos del cumplimiento forzoso de la ley, podría tener sentido trasladar algo de la responsabilidad de la prevención de los delitos a sus víctimas potenciales mediante una reducción de las penas.

táculos para los objetivos de personas dispuestas a matar; para aquellas personas la probabilidad de la muerte sería mayor que la probabilidad de una muerte accidental, haciendo que el otorgamiento de los daños óptimos supere a los de una muerte ordinaria, incluso una muerte causada por negligencia.

También debemos considerar el efecto del ocultamiento sobre el nivel correcto de daños. Los accidentes, que en su mayoría son un subproducto de actividades públicas legales, son de ordinario difíciles de ocultar; y los incumplimientos de contratos son de ordinario imposibles de ocultar. Pero, cuando todo el propósito de alguien es tomar algo de valor de otra persona, tratará naturalmente de ocultar lo que está haciendo y a menudo lo logrará porque ha planeado el ocultamiento con anticipación. La fórmula útil para decidir cuánto otorgar en daños, si la probabilidad de que el cuasidelincuente sea efectivamente atrapado y forzado a pagar los daños es menor que uno, es $D = L/p$, donde D es el otorgamiento de daños óptimos, L es el daño causado por el cuasidelincuente (incluido cualquier ajuste para desalentar la práctica de esquivar al mercado mediante transferencias de riqueza coercitivas), y p es la probabilidad de ser atrapado y obligado a pagar D. Si $p = 1$, L y D son la misma cantidad. Pero si, por ejemplo, $L = 10\,000$ dólares y $p = 0.1$, lo que significa que nueve de cada diez veces el cuasidelincuente escapa a las garras de la ley, entonces D, el castigo óptimo, es 100 000 dólares. Sólo si se aumenta el castigo a este nivel, el costo *esperado* del castigo para el cuasidelincuente potencial (pD) será igual al daño causado por su acto (L).

Una vez que se ajustan hacia arriba los daños en el caso de la transferencia coercitiva pura para desalentar los esfuerzos de hacer a un lado al mercado, para reconocer la relación no lineal existente entre el riesgo de muerte y la compensación por asumir el riesgo y para compensar los esfuerzos del demandado por el ocultamiento, el nivel de los daños será a menudo mayor que la capacidad de pago del cuasidelincuente. Son posibles tres respuestas, todas ellas usadas por la sociedad. Una consiste en imponer la desutilidad al delincuente en formas no monetarias, tales como la prisión o la muerte. Otra consiste en reducir la probabilidad del ocultamiento manteniendo una fuerza policiaca que investigue los delitos. Una tercera, que implica a la vez el mantenimiento de una fuerza policiaca y el castigo de los actos preparatorios (la categoría 4), consiste en impedir la actividad delictiva antes de que ocurra. Si, por las razones que se discutirán en el capítulo XXIII, la vigilancia pública es más eficiente que la privada, el Estado entra al campo del cumplimiento forzoso y tiene un derecho a cualesquiera penas monetarias que se impongan. Así, estas penas se pagan al Estado en forma de multas y no a las víctimas del delito como daños. Las víctimas pueden buscar el pago de daños si el delito es también un cuasidelito, ya sea del derecho común o de las leyes reglamentarias.

Cuando los remedios de los cuasidelitos son un disuasivo adecuado porque los daños, incluidos cualesquiera daños punitivos, se encuentran dentro de la capacidad de pago del demandado potencial, no habrá necesidad de

invocar los castigos penales (los que, como se explicará más adelante, son más costosos que los civiles aun cuando sólo se imponga una multa). Aunque en algunos casos, como en los de antimonopolio y el fraude de valores, los demandados ricos son procesados penalmente y enjuiciados civilmente, las sanciones penales se reservan, como lo pronostica la teoría, para los casos en que el remedio del cuasidelito se estrella contra una limitación de solvencia. Parece que en estos casos el "castigo doble" (cuasidelictivo y penal) de los ricos generaría un castigo excesivo (uno más severo que el óptimo, debido a que cuesta más lograr la severidad agregada, y/o impone una desutilidad mayor al demandado, que el beneficio que produce en cualquier disuasión menor la severidad agregada). Pero su riqueza les permite a los demandados acaudalados contratar buenos abogados, lo que disminuye la probabilidad de ser condenados (aunque sean culpables) y por ello el efecto disuasivo de la amenaza de un castigo penal. Más importante, el demandado rico que "se escabulle" con sólo pagar una demanda por cuasidelito, incluso si la sentencia confisca toda su riqueza existente, puede recuperarse rápidamente porque se mantiene intacto el capital humano (habilidades, contactos, etc.) que generó esa riqueza.

El argumento decisivo para la imposición de sanciones penales a los ricos es que su riqueza, por grande que sea, puede estar por debajo del nivel óptimo del pago de daños debido al ocultamiento del delito. Supongamos que el demandado ha acumulado 100 millones de dólares a través de violaciones penales del derecho bursátil, y la probabilidad de que sea capturado y sometido a un proceso judicial por cuasidelito es de 20%. Por lo tanto, su costo esperado del castigo es de tan sólo 20 millones y no es suficiente para disuadirlo de emprender las actividades criminales que le redituaron 100 millones de dólares. Lo que significa que incluso la perspectiva del otorgamiento de un pago de daños de 100 millones de dólares no sería suficiente para disuadirlo, porque la probabilidad de su imposición es menor a 100 por ciento.

El argumento de que el sistema de derecho penal está primordialmente (aunque no exclusivamente) diseñado para quienes no son ricos no se ve refutado por el uso de multas como un castigo penal. Tales multas son generalmente mucho menores que los daños correspondientes de los cuasidelitos y aun entonces se perdonan a menudo a causa de la indigencia del demandado.

VII.2. Sanciones penales óptimas

Una persona comete un delito porque considera, aunque a menudo de forma errónea o presa de la locura, que los beneficios esperados, ya sea monetarios o emocionales, superan a los costos esperados. Muchos delincuentes tienen un CI bajo o problemas psicológicos (o ambos) que les dificultan obtener un empleo formal o realizar una evaluación competente de los beneficios y los costos relativos (incluidos los costos esperados del castigo) de una actividad

delictiva. No obstante, la literatura empírica sobre el delito ha revelado que los delincuentes responden a los cambios habidos en los costos de oportunidad del tiempo de los delincuentes, en la probabilidad de la aprehensión (y por consiguiente de la cantidad de policías), en la duración de las penas de cárcel y en otras variables relevantes… y esto independientemente de si el delito se comete por la ganancia pecuniaria o por impulso pasional, o por personas instruidas o iletradas,[4] o, para el caso, aun por menores de edad.[5]

Los costos del delito incluyen diversos gastos pecuniarios (para comprar armas de fuego, herramientas de hurto, máscaras, etc.), así como los costos de oportunidad del tiempo del delincuente y los costos esperados del castigo penal. El último de estos costos será el centro de nuestro análisis, pero convendrá mencionar a los otros costos a fin de observar la posibilidad de controlar el nivel de la actividad criminal en otras formas que el jugueteo con el nivel de la actividad de cumplimiento forzoso de la ley y la gravedad del castigo. Por ejemplo, los costos de oportunidad del delito podrían aumentarse, y por ende reducirse la incidencia del delito, disminuyendo el desempleo, lo que aumentaría las ganancias del trabajo legal.[6] Los beneficios del robo, y por ende su incidencia, también podrían disminuirse mediante una redistribución de la riqueza más allá de las manos de los ricos… o aumentarse. Es más fácil cercar los bienes de uso común, y éstos son más ampliamente poseídos en una sociedad igualitaria; y un sistema de bienestar social (un componente tradicional de los programas para la reducción de las desigualdades del ingreso) disminuye los costos de oportunidad del delito al gravar pesadamente al ingreso legítimo. La pérdida de beneficios del bienestar es un costo de la ganancia de dinero suficiente para salir de la beneficencia, y (como un impuesto explícito al ingreso) disminuye el ingreso neto derivado del trabajo.

Aunque la sanción penal pretende disuadir, debemos considerar la posibilidad de que el único efecto del castigo —ya que para la mayoría de los delitos mayores asume la forma de la prisión (en ocasiones, aunque rara vez, la ejecución) antes que de una multa— fuese el de incapacitar a los delincuentes, y de que la mera amenaza de un castigo no tenga ningún efecto disuasivo.

[4] Véanse resúmenes de la literatura empírica sobre el modelo de la elección racional del comportamiento criminal en Steven D. Levitt, "Understanding Why Crime Fell in the 1990s: Four Factors That Explain the Decline and Six That Do Not", *J. Econ. Perspectives*, invierno de 2004, 163; Isaac Ehrlich, "Crime, Punishment, and the Market for Offenses", *J. Econ. Perspectives*, invierno de 1996, 43, 55-63; D. J. Pyle, "The Economic Approach to Crime and Punishment", 6 *J. Interdisciplinary Stud.*, 1, 4-8 (1995). Véase un estudio empírico ilustrativo en Steven D. Levitt, "The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation", 111 *Q. J. Econ.*, 319 (1996).

[5] Véase Steven D. Levitt, "Juvenile Crime and Punishment", 106 *J. Pol. Econ*., 1156 (1998).

[6] Véanse pruebas de que la reducción del nivel de desempleo reduce, en efecto, la incidencia de delitos contra la propiedad, pero que una forma más importante en la que la prosperidad reduce la delincuencia consiste en habilitar presupuestos estatales y locales mayores para la policía y las prisiones, en Levitt, "Understanding Why Crime Fell in the 1990s", nota 4 anterior, en 170-171.

Esto significaría no sólo que los delincuentes son irracionales y por consiguiente incapaces de comprender el significado de un costo en el que sólo tienen la probabilidad de incurrir y no la certeza, sino también que la elasticidad-precio de la oferta de delincuentes (el efecto de un cambio pequeño en el precio sobre la cantidad ofertada) es cero, así que si se saca de la circulación a los delincuentes no se inducirían nuevas entradas al mercado de los delincuentes al aumentar los rendimientos del delito.

La división de los efectos de disuasión y de incapacitación de los castigos es un área empírica desafiante. Un estudio de los programas de castigo de "tres *strikes* y estás fuera" descubre un efecto disuasivo sustancial.[7] En tales programas, una persona que cometa tres delitos será sentenciada a prisión perpetua por el tercero, aunque éste sea relativamente menor. Cuando se promulga la ley, aplicable por supuesto sólo a las personas que cometan un tercer delito después de su fecha de entrada en vigor (porque los castigos penales no pueden aumentarse retroactivamente sin violar la prohibición constitucional de las leyes *ex post facto*), no hay ningún efecto incapacitante inmediato, porque nadie ha cometido todavía el tercer delito que lo sujetaría a la nueva ley. El único efecto inmediato de la ley consiste en aumentar la amenaza de castigo por cometer un tercer delito, lo que significa que si, como los estudios descubrieron que ocurría, hay una disminución *inmediata* de la tasa delincuencial (manteniendo constantes otros factores que influyen sobre ella), ello sólo puede deberse al efecto disuasivo de la ley.

Vimos antes que la sanción penal debe nivelarse de tal modo que el delincuente vea empeorada su situación por la comisión del acto. Pero ahora debemos introducir varias calificaciones. Supongamos que me pierdo en el bosque y, como una alternativa a fallecer de hambre, me meto a una cabaña deshabitada y robo una cantidad trivial de alimentos que encuentro allí. ¿Realmente queremos aplicar la pena de muerte por este robo, según la teoría de que el delito salvó mi vida, de modo que ningún castigo menor sería disuasivo? Por supuesto que no. El problema es que, mientras que la ley del robo castiga generalmente los apoderamientos en contextos de costos de transacción bajos, en este ejemplo son prohibitivos los costos de transacción con el dueño de la cabaña ausente. Un enfoque consistiría en definir el robo de tal modo que se excluyan tales ejemplos; y, en efecto, el derecho penal contiene una defensa de necesidad que probablemente podría invocarse con éxito en este ejemplo (véase después VII.8). Con todo, como veremos en el capítulo XXI, podrían ser muy grandes los costos del intento de hacer una especificación tan detallada del delito, y la alternativa consiste en emplear una definición más general, si bien demasiado incluyente, pero fijando el costo esperado

[7] Véanse Steven D. Levitt, "Why Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error?", 36 *Econ. Inquiry*, 353 (1998); Daniel Kessler y Steven D. Levitt, "Using Sentencing Enhancements to Distinguish between Deterrence and Incapacitation", 42 *J. Law & Econ*., 343 (1999).

del castigo a un nivel que no disuada del delito ocasional que maximiza el valor.

Hay una razón relacionada para imponer un tope a los castigos penales por debajo del nivel en el que todos los delitos sean evitados. Si hay un riesgo de violación accidental del derecho penal (y lo hay, para cualquier delito que implique un elemento de negligencia o de responsabilidad estricta) o de error legal, un castigo muy severo inducirá a las personas a abstenerse de actividades socialmente deseables que se encuentren al borde de la actividad criminal. Por ejemplo, si el castigo por conducir a más de 90 kilómetros por hora fuese la muerte, los individuos conducirían demasiado lentamente (o no lo harían en absoluto) para evitar una violación accidental o una condena por error. Siempre que la categoría de los actos criminales se limite, mediante el concepto de la intencionalidad y defensas tales como la de la necesidad, a casos donde, en los términos de la Fórmula de Hand, hay una disparidad muy grande entre *B* y *PL*, el riesgo del accidente o del error será escaso, y el sistema legal puede sentirse más libre para fijar castigos severos. Pero no totalmente libre; pues, si las consecuencias del error son suficientemente enormes, incluso un riesgo de error muy pequeño hará que generen medidas de evitación que podrían ser muy costosas socialmente. Y, dado que hay costos de inclusión deficiente si los requerimientos de prueba de la culpa se fijan a un nivel muy elevado, podría tener sentido volver más fácil la prueba, pero al mismo tiempo volver menos severo el castigo a fin de reducir los costos de evitar el acto y del error.

¿Cuál es entonces la mejor combinación de probabilidad y severidad del castigo para imponer al delincuente el costo esperado deseado del castigo? Para empezar, supongamos que el castigo toma la forma de multa; a los economistas les agradan las multas como una forma de castigo, ya que son mucho más baratas que la cárcel, dado que la multa es un ingreso para el gobierno y una vez que se ha cobrado el gobierno no tiene más gastos de castigo de la persona multada; eso es muy diferente de tener que pagar los costos de vida de un interno de la prisión.

Un costo esperado del castigo de 1 000 dólares puede imponerse combinando una multa de 1 000 dólares con una probabilidad de aprehensión y condena de 1, una multa de 10 000 dólares con una probabilidad de 0.1, una multa de un millón de dólares con una probabilidad de 0.001, etc. Si los costos del cobro de las multas se suponen iguales a cero cualquiera que sea el monto de la multa, la combinación más eficiente es una probabilidad arbitrariamente cercana a cero y una multa arbitrariamente cercana al infinito. Porque, mientras que los costos de la aprehensión y la condena de los delincuentes aumentan con la probabilidad de la aprehensión —mayores probabilidades implican más policías, fiscales, jueces, abogados defensores, etc., que cuando la probabilidad de la aprehensión es muy baja—,[8] los costos del cobro de las

[8] Sin embargo, sólo la primera implicación es segura: que habrá más policías. Si la disuasión es tan eficaz como lo sería con una combinación entre un castigo más severo y esfuerzos meno-

multas son, según se da por sentado, nulos independientemente de su monto. Así pues, todo aumento del monto de la multa no cuesta nada, mientras que toda disminución correspondiente de la probabilidad de aprehensión y condena, destinada a contrarrestar el aumento de la multa y mantener así un costo constante del castigo esperado, disminuye los costos del cumplimiento forzoso, hasta el punto del desvanecimiento si la probabilidad de la aprehensión y la condena se reduce a un nivel arbitrariamente cercano a cero.

Pero hay varios problemas con el supuesto de que el costo de la multa no se relaciona con su monto. Primero, algunos criminales sienten aversión por el riesgo. La razón de que en el modelo de castigo mediante multa el único costo de una multa sea el costo de su cobro y no el monto en dólares de la multa misma, es que la multa no se paga (porque se impide el delito) o, si se paga, simplemente transfiere una cantidad de dólares igual del criminal al contribuyente. Pero en el caso de los criminales que sienten aversión por el riesgo, toda disminución de la probabilidad de la aprehensión y la condena, y todo aumento correspondiente de la multa para quienes son aprehendidos y condenados, imponen una desutilidad que no se traduce en recaudación para el Estado.[9] Y tampoco se contrarresta este efecto por el efecto sobre los criminales que prefieren el riesgo, aunque su número sea igual al de los criminales que sienten aversión por el riesgo, suponiendo que no se pueden distinguir las dos clases en el proceso de sentencia. En la medida en que una multa mayor con menor probabilidad de imposición aumente la utilidad de quienes prefieren el riesgo, la multa tendrá que aumentarse un poco para asegurarse de que los disuada, lo que vuelve las cosas más dolorosas aun para quienes le tienen aversión.

El efecto de estigma de una multa (como el de cualquier castigo penal) tampoco se transfiere al Estado. Al igual que la desutilidad impuesta a los delincuentes con aversión al riesgo, se trata de una pérdida social de peso muerto, aunque, en la medida en que el estigma daña al delincuente sólo porque transmite información útil a los potenciales realizadores de transacciones con el criminal condenado (véase la discusión del derecho a la privacidad después en XI.4), crea un valor social que podría contrarrestar el daño.

Otro límite a la severidad de las multas es que el costo de la colecta aumenta con el tamaño de la multa, y otro más es que sin un tope todos los delitos se castigarían con una multa uniformemente severa, ya que, para cada delito, considerado de manera aislada, la multa óptima sería, según nuestro análi-

res para imponerlo, es posible que no haya más investigaciones o juicios penales de los que habría bajo el régimen mencionado en último término. En efecto, con una disuasión completa no habría investigaciones ni juicios bajo ninguno de los dos regímenes, porque no se cometería ningún delito.

[9] Esto supone que el bienestar de los criminales es parte del bienestar social. ¿Está justificada la suposición? Ésa es una cuestión ética. Desde el punto de vista de la maximización de la riqueza, la suposición es correcta.

sis de este punto, cercana al infinito a fin de minimizar el costo de la aprehensión y la condena de los delincuentes. Pero el castigo uniforme de los delitos de gravedad distinta es incompatible con la disuasión marginal: la creación de incentivos para que los delincuentes incorregibles sustituyan los delitos más graves por otros menos graves. Si el robo se castiga tan severamente como el homicidio, el ladrón podría igualmente asesinar a su víctima para eliminar a un testigo, porque esto reduciría la probabilidad de que fuera aprehendido sin aumentar el castigo en caso de que eso ocurriera, y reduciría así pD. Puede preservarse la disuasión marginal variando la probabilidad del castigo con la gravedad del delito, es decir, buscando más asiduamente a los ofensores más graves. Esto se hace en una medida considerable, pero aumenta el costo del sistema de justicia penal.

Con todo, la eliminación de la disuasión marginal, al menos en algunas clases de caso, podría no ser una mala idea del todo. Vuélvase punible al robo con la pena de muerte y, en efecto, aumentará la probabilidad de que *si* se cometiera un robo alguien sería asesinado en el proceso, pero disminuirá la probabilidad de que el robo se cometiera en primer lugar. El número total de los homicidios cometidos durante los robos podría por ello disminuir (porque habrá muchos menos robos), aunque fuese mayor el incentivo de los ladrones para asesinar. Aun así, la eliminación de la disuasión marginal de un delito aumentará indudablemente la incidencia del delito si la ofensa menor que se castigará ahora con igual severidad es un sustituto, antes que un complemento, de la ofensa mayor. El robo y el homicidio en el curso del robo son complementos, porque, siendo todas las demás consideraciones iguales, un aumento del primero conducirá también a un aumento del segundo. Pero, por ejemplo, si se aumentara el castigo para el robo de bicicletas al mismo nivel que para el robo de automóviles, aumentaría la incidencia del robo de automóviles, ya que todo lo que habría cambiado sería que el precio relativo de un sustituto aumentó. Los ladrones de bicicletas comenzarían a robar automóviles en lugar de bicicletas.

Las multas no son una panacea del castigo. Porque la mayoría de los criminales no tienen los medios para pagar multas fuertes, los sistemas de justicia penal dependen mucho, por fuerza, de las sanciones no pecuniarias. La prisión sí impone costos pecuniarios y no pecuniarios al infractor, al reducir su ingreso durante y generalmente después del periodo de confinamiento. Pero ésta es una razón más para preferir el encarcelamiento a la multa como la sanción común de la conducta delictiva. El encarcelamiento reduce la riqueza incluso de los criminales (con mucho, la gran mayoría) que no pueden pagar ninguna multa.

Dado que las multas y la prisión son simplemente maneras diferentes de imponer desutilidad a los infractores, la Corte Suprema se equivoca cuando considera que una sentencia que impone una multa pero provee el encarcelamiento si el demandado no puede pagarla es discriminatoria contra los po-

bres.[10] Puede encontrarse una tasa de cambio que iguale, para un individuo dado, un número de dólares con un número de días en la cárcel. Pero quizá la Corte objete realmente el hecho de que la mayoría de las leyes penales establecen una tasa de cambio muy favorable para las personas que tienen dinero. Quinientos dólares es una sanción más moderada que la de 100 días en la cárcel (Williams *vs*. Illinois), aun para las personas de ingresos bajos; ésta es una sanción trivial para otras personas: las que tienen mayores probabilidades de ser capaces de pagar la multa en lugar de pasar un tiempo en la cárcel.

Los costos sociales de la prisión superan enormemente a los del cobro de multas a los demandados solventes. Existe el gasto de la construcción, el mantenimiento y la operación de las prisiones, la pérdida de la producción legítima del individuo encarcelado (si hay alguna) mientras se encuentre en prisión, la desutilidad de la prisión para él (lo que no genera ningún beneficio correspondiente para el Estado, como lo hace una multa), y el probable deterioro de su productividad en actividades legítimas después de la liberación. El deterioro que es relevante aquí no es el causado por el estigma de la condena, que es independiente de la forma específica del castigo (aunque no de su severidad): es aquel causado por la depreciación de las habilidades, la pérdida de contactos, etc., durante el periodo del encarcelamiento; en suma, la depreciación del capital humano del condenado. Dado que el ingreso perdido del empleo legítimo es un costo de oportunidad del delito, una disminución de las perspectivas de ingresos legítimos del prisionero disminuye los costos de la actividad criminal para él y, por ende, aumenta la probabilidad de que cometa delitos después de su liberación. Pero el encarcelamiento genera un beneficio importante que una multa no genera: impide que el delincuente cometa delitos (¡por lo menos fuera de la prisión!) mientras permanezca en prisión.

Es mucho lo que puede hacerse para mejorar los castigos alternativos al encarcelamiento. Las multas pueden hacerse pagaderas en abonos. Pueden hacerse proporcionales a los ingresos y pagaderas con ellos, en lugar de ser una cantidad fija en dólares. Puede usarse como una sanción la exclusión de ocupaciones particulares. La libertad de acción puede restringirse (y con frecuencia se restringe ahora) en formas que permitan la actividad remunerativa, como ocurre con el encarcelamiento sólo por la noche y los fines de semana, la cual permitiría que el criminal pagara una multa. Pero algunos de estos métodos no están enteramente libres de las desventajas del encarcelamiento. Una multa pagadera en abonos o proporcional a los ingresos futuros reduciría el ingreso del infractor derivado de la actividad lícita y, por ende,

[10] Véase Tate *vs.* Short, 401 U.S., 395 (1971); Williams *vs.* Illinois, 399 U.S., 235 (1970). Sin embargo, más tarde, la Corte Suprema señaló que, "si las medidas alternativas no son adecuadas para satisfacer los intereses del Estado relativos al castigo y la disuasión, [el tribunal puede] encarcelar a una persona en libertad condicional que ha hecho suficientes esfuerzos de buena fe para pagar". Bearden *vs*. Georgia, 461 U.S., 660, 672 (1983).

también su incentivo para preferirla a la actividad criminal, como ocurriría también con la exclusión de una ocupación.

Si las multas son en efecto un método de castigo más eficiente que la prisión cuando el delincuente es solvente, ¿cómo explicar el uso creciente de las sentencias de prisión para castigar los delitos típicos de los ricos, tales como la fijación de precios, la evasión fiscal, el fraude con valores y el soborno? La interrogante no es idéntica a la que busca la razón por la que se procesa en cualquier caso a los delincuentes ricos, antes que dejárselo al sistema de los cuasidelitos, porque una condena penal conlleva un estigma, incluso cuando el único castigo es una multa, que puede perjudicar el capital humano del demandado, así como desencadenar consecuencias colaterales como ser excluido de un empleo futuro en la industria de valores, si se trata de una infracción penal de la ley bursátil. Entre mayor sea el capital humano que tiene el demandado, mayor será el deterioro del mismo derivado de la condena por un delito.[11] Sin embargo, la limitación de la solvencia que discutimos anteriormente continúa siendo un argumento decisivo a favor de la pena no pecuniaria, aunque si las preocupaciones por la eficiencia en verdad figuran en la práctica de las sentencias esperaremos (y descubrimos[12]) sentencias de multas más fuertes y periodos de prisión más cortos para los acusados más ricos (manteniendo constantes otras características del delito y del delincuente) porque el encarcelamiento es más costoso que una multa para el Estado, cuando la multa es cobrable, y también porque el encarcelamiento impone una mayor desutilidad entre mayores sean las diferencias entre el nivel de vida del prisionero antes de que se le encarcelara y su nivel de vida en la prisión.

El costo del cobro de las multas y el del encarcelamiento han revivido el interés por los castigos "vergonzosos",[13] que antes eran comunes (por ejemplo, los cepos), y que son más baratos que la prisión y, no obstante, no tienen los problemas de liquidez que frustran el cobro de las multas elevadas. Una pena vergonzosa pura inflige desutilidad bajo la forma de la vergüenza o la humillación, pero a menudo también provee de información útil, por ejemplo, cuando se requiere que una persona que haya sido encontrada culpable de prostituir mujeres publique un aviso en ese sentido en el frente de su jardín o en la ventanilla trasera de su automóvil. Sin embargo, transmitir esa información (hoy en día, el ejemplo más importante es que los convictos por

[11] Véanse pruebas en John R. Lott Jr., "An Attempt at Measuring the Total Monetary Penalty from Drug Convictions: The Importance of an Individual's Reputation", 21 *J. Legal Stud.,* 159 (1992).

[12] Véase Joel Waldfogel, "Are Fines and Prison Terms Used Efficiently? Evidence on Federal Fraud Offenders", 38 *J. Law & Econ.*, 107 (1995). Véase una discusión más amplia en la siguiente sección.

[13] Véase, por ejemplo, Dan Kahan y Eric A. Posner, "Shaming White-Collar Criminals: A Proposal for Reform of the Federal Sentencing Guidelines", 42 *J. Law & Econ*., 365 (1999).

abusar sexualmente de menores de edad sean registrados en bases de datos públicas) podría reducir las oportunidades de empleos legítimos del delincuente más aún de lo que lo haría su relación de antecedentes penales; en tal caso, disminuirá para él el costo de oportunidad de cometer delitos en el futuro. Además, si el castigo y la vergüenza son complementos en el sentido de que entre más severo sea el castigo más vergüenza se añadirá al acto criminal, disminuir la multa o la duración del encarcelamiento por un delito mediante la sustitución parcial de una pena vergonzosa podría volver al castigo menos eficaz al disminuir lo vergonzoso del delito, mientras que si el castigo vergonzoso simplemente se suma a una multa o un encarcelamiento invariables no habrá ningún ahorro de costos para el gobierno.

Aun un encarcelamiento prolongado podría no imponer a un homicida costos iguales a los de la víctima, pero esto no es en sí mismo una justificación económica para la cada vez más cuestionada pena capital en el caso del homicidio. Lo importante no es que el castigo del homicidio sea proporcional al costo de la víctima, sino que sea suficientemente severo para que el homicidio no sea rentable… y seguramente la cadena perpetua le costaría al asesino más de lo que posiblemente hubiera ganado, con muy raras excepciones. Pero, al igual que en el caso de la multa para los delincuentes ricos, este análisis trata implícitamente la probabilidad de la aprehensión y de la condena como igual a uno. Si es menor que uno, como lo es por supuesto, el asesino no estará comparando la ganancia del delito con el costo si es atrapado y condenado: estará comparándola con el costo de la condena descontado por la probabilidad de que sea atrapado y condenado.

El argumento en favor de la pena capital por motivos de disuasión es particularmente sólido cuando los delincuentes tienen tasas de descuento muy altas (véase la siguiente sección de este capítulo), por lo que son relativamente insensibles a la duración de una condena de prisión, o enfrentan riesgos de muerte altos por otras causas y, en consecuencia, tienen una esperanza de vida corta. Pero el argumento no es concluyente. En virtud de que el costo de la imposición errada es tan severo, se invierten recursos sustancialmente mayores en la litigación de un caso de pena capital (véase después XXIII.4) y se tiene que compensar el costo adicional con cualquier disuasión adicional, la cual puede ser menor. Porque debemos distinguir entre la pena capital como un castigo común para los homicidas y la pena capital como se administra actualmente en los Estados Unidos. El sentido común sugiere que, si se impone habitualmente, la pena capital disuadiría efectivamente a las personas por el miedo a la muerte; después de todo, es sumamente raro que un preso condenado a muerte rechace la conmutación de su condena a cadena perpetua sin posibilidad de libertad condicional. Pero, en parte debido a las restricciones procedimentales impuestas por la Corte Suprema, la pena capital se impone tan pocas veces que es poco probable que su expectativa logre la disuasión. "A pesar de los aumentos de la pena capital en los últimos años,

la probabilidad de ser ejecutado condicionada a la comisión de un asesinato todavía es menor a 1 por cada 200. Incluso entre los condenados a muerte, el porcentaje de ejecución anual es de sólo 2%, o [sólo] el doble de la tasa de mortalidad por accidentes y violencia entre los hombres estadunidenses."[14]

Sin embargo, en casos especiales la pena capital es apoyada fuertemente también por consideraciones de la disuasión marginal. Si la pena máxima por el homicidio es la cadena perpetua, quizá no deseemos castigar también el robo a mano armada con la cadena perpetua. Pero si así disminuimos el castigo máximo para el robo armado de la cadena perpetua a 20 años de prisión, no podremos castigar con 20 años algún delito menor. Empero, no se sigue de aquí que la pena capital debiera ser el castigo para el homicidio *simple*. Porque entonces tendremos el problema de la disuasión marginal del homicida múltiple. Es posible que la pena capital debiera estar reservada para él, de modo que los homicidas tengan un desincentivo para asesinar a los testigos del homicidio. Una aplicación importante de este punto ocurre en los asesinatos en prisión. Si un prisionero ha sido condenado a cadena perpetua por homicidio, no tendrá ningún desincentivo para no asesinar en la prisión a menos que el homicidio en la prisión pueda castigarse con la muerte.[15]

Problemas de esta clase mortificaban a los pensadores medievales. En virtud de que la mayoría de los hombres medievales creían en otra vida y de que la expectativa de vida era corta, la pena capital no era un castigo tan grave en esos días como en nuestro mundo moderno y (según parecía hasta hace poco tiempo) cada vez más secular. En un intento por volver más cara la pena capital, se prescribían métodos de ejecución horribles (como el arrastramiento y el descuartizamiento) para delitos particularmente aborrecibles como la traición. El hervido en aceite, considerado más horrible que el ahorcamiento o la decapitación, se empleaba para castigar el homicidio por envenenamiento; dado que los envenenadores eran especialmente difíciles de aprehender en esa época, un castigo más severo que el prescrito para los homicidios ordinarios estaba (económicamente) indicado. Otro ejemplo en defensa del punto general —que entre más escasa sea la probabilidad del castigo más severo debe ser éste para preservar la disuasión— era el ahorcamiento de los ladrones de caballos en el Oeste estadunidense del siglo XIX. Otro más era la pena de muerte (que no siempre se ejecutaba, sobre todo en el caso de quienes delinquían por primera vez) para todos los delitos graves, y muchos de los no tan graves, en el siglo XVIII en Inglaterra, donde no había ninguna fuerza policiaca organizada y la probabilidad del castigo era en consecuencia muy baja. Una consideración adicional era que la pena de muerte solía ser mucho más barata que el encarcelamiento. Los costos procedimentales de la pena capital han aumentado enormemente en los Estados Unidos, al igual

[14] Levitt, "Understanding Why Crime Fell in the 1990s", nota 4 anterior, en 175.

[15] O a menos que las condiciones de la prisión varíen sustancialmente en su severidad, reservándose las peores para quienes asesinen en la prisión. Esto se hace en alguna medida.

que el intervalo entre el juicio y la ejecución (consecuencia de la disposición de remedios elaborados después de la condena), como resultado de lo cual es probable que la persona condenada se encuentre en el corredor de la muerte al menos durante una década antes de su ejecución. En efecto, el costo de la pena capital es aditivo a un costo sustancial de la prisión.

Incluso si la pena capital por el homicidio no tiene un efecto de disuasión incremental, confiere utilidad a la familia y los amigos de la víctima. Desde un punto de vista económico, no hay ninguna razón para no incluir esa utilidad en la decisión de conservar o no la pena capital. Se puede argumentar en contra de esto que la familia de la víctima y sus amigos derivan esta utilidad adicional del hecho de que la pena capital impone una mayor desutilidad al homicida que la que impone el encarcelamiento (en otras palabras, se trata de un caso de utilidades negativas interdependientes, como se discute antes en VI.15), por lo que puede no haber una ganancia neta en el bienestar social.

Dado que la ley debe continuar utilizando en gran medida el encarcelamiento como una sanción penal, hay un argumento —sujeto a reservas que ya deben ser familiares para el lector, basadas en la aversión al riesgo, la inclusión excesiva, los costos de evitar el acto y el error y (posiblemente) la disuasión marginal— a favor de que se combinen los periodos de prisión extensos para los delincuentes condenados con bajas probabilidades de aprehensión y condena. Consideremos la elección entre la combinación de una probabilidad de aprehensión y condena de 0.1 con una prisión de 10 años con una combinación de una probabilidad de aprehensión y condena de 0.2 y una prisión de cinco años. Bajo el segundo enfoque se envía a prisión al doble de individuos, pero sólo por la mitad de la duración de la prisión. Por lo tanto, los costos totales del encarcelamiento son similares bajo los dos enfoques; pero los costos de la policía, los funcionarios judiciales, etc., son menores con el primer enfoque. ¿Pero no es injusto un sistema donde las probabilidades del castigo son bajas, porque crea una desigualdad *ex post* entre los infractores? Muchos se escapan; otros pasan más años en prisión de lo que ocurriría si se atrapara a más infractores. Sin embargo, objetar este resultado equivale a decir que todas las loterías son injustas porque crean, *ex post*, diferencias de riqueza entre los jugadores. En un sentido igualmente significativo, tanto el sistema de justicia penal que crea bajas probabilidades de aprehensión y condena como la lotería son justos siempre que se igualen entre los participantes los costos y los beneficios *ex ante*.

Ahora bien, una de las salvedades requiere énfasis. Quienes prefieren el riesgo prefieren un enfoque de justicia penal que combine condenas fuertes con bajas probabilidades de imposición. Si, como parece creíble, los delincuentes son más propensos a preferir el riesgo que a sentir aversión por él, dado que el delito es una tarea inherentemente riesgosa, el enfoque de alta probabilidad de aprehensión-condena menor puede ser superior. Otro problema con el enfoque de baja probabilidad-gran severidad de la condena surge del hecho de que un periodo de prisión se alarga, por supuesto, añadiendo

tiempo al final. A una tasa de descuento de 10%, un confinamiento en la prisión de 10 años impone una desutilidad sólo 6.1 veces mayor que la desutilidad de una sentencia de un año, y una sentencia de 20 años aumenta esta cifra a sólo 8.5 (las cifras correspondientes para una tasa de descuento de 5% son 7.7 y 13.5). Los criminales tienden a tener tasas de descuento altas, debido a que una tasa de esta naturaleza llevará a una persona a invertir menos en educación y, por tanto, a reducir sus oportunidades de tener ingresos legítimos y, por consiguiente, a tener costos de oportunidad del delito más bajos. Incidentalmente, éste es un argumento a favor de los términos de prescripción más cortos en los casos penales que en los casos civiles. El plazo de prescripción corto ejerce presión sobre la policía para atrapar a los delincuentes más temprano que tarde, lo que favorece una mayor disuasión si la mayoría de los delincuentes tienen tasas de descuento altas.[16]

Otro argumento a favor del enfoque de alta probabilidad de aprehensión-condena menor es que, si la gente no tiene buena información acerca de la probabilidad de la aprehensión y la condena y, en consecuencia, algunos delincuentes potenciales sobrestiman tal probabilidad, mientras que otros la subestiman, tanto la disuasión excesiva como la disuasión deficiente podrán disminuirse aumentando la probabilidad y disminuyendo el castigo correspondientemente. Supongamos que f (el castigo) es 500 con $p = 0.02$ (de modo que el costo esperado del castigo, fp, es 10), y la mitad de la población relevante sobrestima p en 1%, calculando así p como 0.03, de tal modo que para ellos la sanción esperada es 15, y la otra mitad subestima p en 1%, calculando así p como sólo 0.01, lo que implica una sanción esperada de 5. Habrá un exceso y una deficiencia de la disuasión. Si en cambio se fija f en 100 y p en 0.1 (generando el mismo costo esperado del castigo, o sea 10), la mitad de la población que sobrestima calculará la sanción esperada en 11 ($100 \times (0.10 + 0.01)$), y la mitad de la población que subestima la estimará en 9; ambas cifras son muy cercanas al costo efectivamente esperado del castigo de 10.[17]

VII.3. Prevención del delito: leyes de infractores múltiples, tentativa y conspiración, instigación y complicidad, reincidencia, entrampamiento

En el modelo económico más simple de la sanción penal, el Estado reduce la demanda del delito fijándole un "precio" bajo la forma de un costo esperado del hecho de tener que pagar una multa o ir a prisión sólo para mantener la credibilidad de la disuasión. Sin embargo, esta concepción deja sin explicar muchos aspectos importantes del sistema de la justicia penal. Por ejemplo,

[16] Yair Listokin. "Efficient Time Bars: A New Rationale for the Existence of Statutes of Limitations in Criminal Law", 31 *J. Legal Stud.*, 99 (2002).

[17] Véase Lucian Arye Bebchuk y Louis Kaplow, "Optimal Sanctions When Individuals Are Imperfectly Informed About the Probability of Apprehension", 21 *J. Legal Stud*., 365 (1992).

un criminal condenado que ha sido sentenciado por delitos anteriores será castigado de ordinario con mayor severidad que quien cometa un delito por primera vez, aun cuando haya cumplido plenamente cualesquiera sentencias que se le impusieran por los delitos anteriores. En los mercados competitivos no se cobra a los consumidores precios mayores sólo porque hayan comprado el mismo producto antes. Ciertamente no se les exige que devuelvan la cosa que hayan comprado si no la han consumido todavía, como se le exigiría a un ladrón. Desde este punto de vista resulta también desconcertante el castigo de los llamados delitos incoados, como las tentativas y las conspiraciones (frustradas). Si el propósito del derecho penal es el de obligar al delincuente a tomar en cuenta plenamente los costos de sus actos, ¿por qué debiera castigársele cuando su conducta no impone ningún costo porque se ha frustrado? ¿Y por qué se considera un valor adicional a favor del encarcelamiento que impida nuevos actos criminales por parte del delincuente encarcelado: los actos que cometería si no estuviera en prisión? Si el sistema de la justicia penal mantiene una lista apropiada de precios para los actos ilícitos, ¿por qué debiera preocuparle a alguien que el delincuente, si no fuera encarcelado sino castigado con severidad equivalente por algún método que lo dejara en libertad, pudiera cometer nuevos actos criminales? Presumiblemente, sólo lo haría si los actos estuviesen justificados por su costo social.

Para contestar estas interrogantes, debemos empezar por señalar que el hincapié en la prevención del delito, antes que en la simple fijación de un precio para él, se aplica a los delitos del derecho común, cuya esencia es una transferencia forzosa en un contexto de costos de transacción bajos. Muy poco de la actividad criminal de esta categoría se justifica socialmente por su costo; los ejemplos como el del robo en una cabaña en condiciones de gran necesidad son raros, y es posible que no sean delitos por la defensa de la necesidad. Así pues, la elevada incidencia de los delitos del derecho común no refleja su deseabilidad social (cercana a cero[18]) sino las dificultades de ubicar el castigo a un nivel suficientemente elevado para lograr una disuasión de 100%. Si, de no ser por el alto costo de las sanciones penales, el nivel óptimo de la actividad criminal sería cero o cercano a cero, estas sanciones no son realmente precios destinados a racionar la actividad; su propósito es, en la medida en que lo permitan los costos, extirparla.[19]

[18] Pues téngase presente que no puede decirse acerca de un delito cometido por alguien imposibilitado para pagar una condena monetaria igual a los costos sociales de su delito que se justifique socialmente por su costo, aunque el delincuente esté dispuesto a incurrir en costos no pecuniarios de la prisión que sean mayores a tales costos sociales. El concepto económico del valor se basa en la disposición a dar algo de valor a otros —pagar—, antes que en la disposición a sufrir una privación que no confiere ningún beneficio a nadie más. El sufrimiento no es un acto productivo que establezca un derecho económico sobre los escasos recursos de la sociedad. Ésta es una diferencia entre un enfoque económico y uno estrictamente utilitario del valor.

[19] Una distinción destacada por Robert Cooter, "Prices and Sanctions", 84 *Colum. L. Rev.,* 1523 (1984).

Este análisis explica el hincapié del derecho penal en la prevención, lo que no tendría sentido en un contexto de mercado o aun en un caso de cuasidelito no intencional. Esto explica la razón por la que un ladrón que es atrapado debe entregar lo que haya robado, aparte de cualquier otro castigo que se le imponga, aunque la víctima no esté buscando la restitución (¡es posible que la víctima sea otro ladrón!). Y ésta es una razón adicional para que las multas deban ser proporcionales a la riqueza del delincuente, aparte de cualesquiera nociones de una distribución justa de la riqueza. El principio de la utilidad marginal decreciente del ingreso implica que se requiere una multa mayor para imponer la misma desutilidad a una persona rica que a una pobre, siempre que ricos y pobres tengan en promedio las mismas funciones de utilidad marginal. (Por lo que toca a la significación de esta calificación, véase después XVII.2.) Lo importante no es que la igualación de las desutilidades sea un objetivo en sí misma, sino que la disuasión óptima pueda requerir en efecto la discriminación de precios en contra de los ricos.

Asimismo, la práctica, sistematizada en las leyes para los infractores múltiples y en las instrucciones para las sentencias federales, de castigar a los delincuentes reincidentes con mayor severidad que a quienes delinquen por primera vez, eleva el precio del delito para quienes, a juzgar por su comportamiento anterior, valúan el delito más que otras personas. Dado que nuestro objetivo es la minimización de la cantidad del delito, se cobra más a quienes valúan más la actividad para participar en ella.

Aún hay más razones para el castigo más severo de los infractores reincidentes que para los que cometen un delito por primera vez:

1. El efecto de estigma del castigo criminal podría disminuir con los castigos sucesivos; es probable que el efecto de los antecedentes penales sobre las oportunidades de empleo, que es una parte importante del efecto de estigma, sea total o casi total después de la primera condena.
2. El hecho de que el acusado haya cometido delitos anteriores aumenta nuestra confianza en que es realmente culpable del delito que ahora se le imputa; por lo tanto, es menor el riesgo de error si se impone una sentencia severa.
3. El castigo más severo de los transgresores reincidentes contrarresta un efecto de "aprendizaje por la experiencia": los transgresores se volverán más hábiles en la comisión de delitos y, por ende, con menores probabilidades de ser detectados a medida que adquieren mayor experiencia. En consecuencia, p bajará, requiriendo un aumento de D aunque L permanezca constante (véase antes VII.1). Sin embargo, puede argüirse que los reincidentes que son atrapados son los que tienen menos probabilidades de eludir a la justicia.
4. El castigo mayor para los reincidentes contrarresta un problema del "último periodo". Entre más viejo sea el delincuente, más difícil será

disuadirlo mediante la amenaza de la prisión, porque la duración del máximo periodo de prisión será truncada por su esperanza de vida disminuida. El castigo más severo para los reincidentes opera para eliminar al criminal de carrera de la población criminal antes de que llegue a la edad en la que la amenaza de la prisión tendrá un efecto de disuasión marcadamente disminuido.

La premisa de la teoría preventiva del castigo penal es que la elasticidad de la oferta de infractores no es infinita. Si lo fuera, esto significaría que un aumento pequeño del rendimiento esperado de la actividad criminal generaría un aumento enorme (literalmente, infinito) de la oferta de infractores, ya que algunas personas en actividades legítimas confluirían en actividades criminales ahora que estas últimas son más rentables; en tales circunstancias, el encarcelamiento de un infractor, que crearía oportunidades para otro, simplemente atraería a una persona de la actividad legítima a la criminal o haría que un criminal de tiempo parcial cometiera más actos criminales. La elasticidad de la oferta de delitos adquisitivos podría ser en efecto muy elevada, pero no es infinita, y presumiblemente es menor la elasticidad de la oferta de delitos pasionales. La avaricia es una característica humana universal, por lo que se puede esperar que un aumento en el ingreso esperado derivado de la actividad delictiva atraiga nuevos participantes al mercado. Los crímenes pasionales tienden a ser una situación específica; el hecho de que en un año menos personas mataran a sus esposas no tendría mucho efecto en la decisión del esposo de X de matar o no a su esposa.

Consideremos ahora el castigo de las tentativas. Un hombre entra a un banco con la intención de robarlo, pero un guardia lo detecta y lo aprehende antes de que pueda hacer algún daño. El hecho de que hubiera llegado tan cerca del robo al banco indica que es muy probable que lo intente de nuevo a menos que se le restrinja, de modo que poniéndolo en prisión probablemente podremos impedir algunos robos. De igual modo, al castigar la tentativa se incrementan los costos esperados del robo de banco para el ladrón sin volver más severo el castigo del robo de banco. El delincuente no puede estar seguro de que su tentativa será exitosa, y si fracasa no sólo perderá las ganancias de un robo exitoso, sino que incurrirá en costos (castigos) adicionales. Así pues, el castigo de las tentativas es como el mantenimiento de una fuerza policiaca: Se incrementa el costo del castigo esperado por el delito consumado sin aumentar la severidad del castigo de ese delito, y a menudo ésta es la combinación eficiente.

La tentativa no se castigará tan severamente como el delito consumado, y hay dos razones económicas para ello: *1)* otorgar al infractor un incentivo para que cambie de opinión en el último minuto (una forma de la disuasión marginal) y *2)* minimizar los costos del error, ya que hay una probabilidad mayor de que el acusado sea realmente inocuo que la de que lo sea alguien que

consumó el delito. Pero el hecho de que la tentativa causara menos daño que el que habría causado el delito consumado no es una razón económica suficiente para el castigo menos severo de la tentativa. El propósito de castigar el intento es prevenir un delito futuro, el delito consumado hacia el que el intento fue un esfuerzo infructuoso. Por lo tanto, el costo relevante es el costo que impondría ese delito antes que el costo que impondría el intento, que de hecho puede ser cero.

¿Qué diríamos si el acusado le hubiese dicho simplemente a un amigo (que resultó ser un informante de la policía): "Me propongo robar ese banco", y no hubiera dado ningún paso para la consumación de su objetivo? Esto no se considerará una tentativa. La probabilidad de que en efecto hubiera robado un banco es menor en este caso que cuando es atrapado a punto de hacerlo, de modo que los beneficios sociales de su encarcelamiento son menores; o, dicho de otro modo, los costos esperados del error son mayores. Desde luego, cuanto más grave sea el delito, más fuerte será el caso para castigar la actividad preparatoria remota; en el cálculo del daño esperado, *PL*, una reducción de *P* puede compensarse por un aumento de *L*. Luego de que los ataques terroristas del 11 de septiembre de 2001 demostraran que el terrorismo era una amenaza mayor para los Estados Unidos de lo que se creía, el Congreso hizo del "material de apoyo" (en un sentido amplio) al terrorismo un delito. La Ley Smith, que penalizaba la conspiración para abogar por el derrocamiento violento del gobierno, se basó en una conceptualización similar.

Las tentativas fallan a veces, no porque sean interrumpidas, sino porque el infractor comete un error. Es posible que le haya disparado a lo que pensó que se trataba de un hombre durmiendo en una cama, pero resultó ser una almohada. O quizás haya hecho un muñeco de vudú de su enemigo, apuñalándolo repetidamente, en la creencia errada de que así mataría al enemigo. La pregunta para el economista es si la naturaleza del error es tal que resulta improbable que el responsable de la tentativa tenga éxito jamás. Si es así no se prevendrá ningún delito enviándolo a prisión. Pero, aun en el caso del muñeco de vudú, si suponemos que el actor aprenderá de la experiencia y así cambiará a un método más eficaz la próxima vez, el hecho de que el primer intento no pudiera haber tenido éxito jamás no sería una buena razón para dejarlo en libertad.[20]

La tentativa que fracasa debido a un error y no porque haya sido interrumpida provee el argumento más fuerte para castigar una tentativa con menor severidad que al delito consumado. Si el castigo para la tentativa de homicidio fuese el mismo que para el homicidio, alguien que le dispare a su pretendida víctima pero falle (y no sea aprehendido de inmediato) podría

[20] Véanse más detalles sobre la economía de las tentativas imposibles en Steven Shavell, "Deterrence and the Punishment of Attempts", 19 *J. Legal Stud.*, 435 (1990); David D. Friedman, "Impossibility, Subjective Probability, and Punishment for Attempts", 20 *J. Legal Stud.*, 179 (1991).

intentarlo de nuevo, porque si logra su propósito será castigado con severidad no mayor que la aplicada a su tentativa infructuosa. Éste es un ejemplo dramático de la importancia de la disuasión marginal. Pero, a fin de que funcione la disuasión en estas situaciones, la utilidad para el infractor, derivada de la comisión del delito consumado, deberá bajar después de la tentativa. Si la utilidad es la misma después de la tentativa lo intentará de nuevo. Sin embargo, a menudo disminuirá porque la pretendida víctima tomará mayores precauciones (si sabe que se hizo el intento, incluso si no sabe quién lo hizo) o porque cambiará de parecer quien llevara a cabo la tentativa, lo que es particularmente probable si la tentativa fue impulsiva antes que deliberada.

Una conspiración que no triunfa en la comisión de un acto criminal es una forma de tentativa, pero a diferencia de ésta es punible aun cuando los conspiradores no se acerquen en ninguna medida a la escena del delito, sino que sean aprehendidos en la etapa preparatoria más temprana de su proyecto. Si las conspiraciones son, en promedio, más peligrosas que los delitos cometidos por un solo hombre, el daño esperado puede ser tan grande como en el caso de la tentativa de un solo hombre aun cuando la probabilidad de la consumación del delito sea menor debido a que los preparativos se interrumpen más temprano. Las conspiraciones son más peligrosas que los delitos cometidos por un solo individuo.[21] Si fuesen más peligrosas sólo en el sentido de que sirven de base a la comisión de más delitos, o delitos más graves, no habría necesidad de un castigo extra; de todos modos, el castigo sería mayor. Pero las conspiraciones son también más peligrosas porque connotan la capacidad de cometer delitos más eficientemente (en un sentido privado, no social) ya que pueden aprovechar la división del trabajo, por ejemplo, apostando a un hombre como centinela, mientras que otro conducirá el automóvil de la fuga, otro más ocultará los bienes robados, etc. Por lo tanto, sus costos son menores (una conspiración simula un enfoque de mercado sobre la comisión del delito) y también es probable que sean más eficaces en la consumación del delito y en evitar la aprehensión; por ambos motivos, el castigo óptimo es mayor. También es cierto que una conspiración es más vulnerable a la detección debido a la escala de sus actividades y (lo que se relaciona directamente) al número de participantes, algunos de los cuales podrían arrepentirse e informar al gobierno. Pero la escala podría permitir también que la conspiración escapara al castigo corrompiendo a los funcionarios encargados de la aplicación de la ley. Y algunos de los delitos más graves, tales como la insurrección, sólo pueden ser cometidos por conspiraciones.

[21] Lo cual es la razón por la que una venta ilegal no es en el derecho una conspiración entre el vendedor y el comprador, y por la que un soborno no es una conspiración entre la persona que paga y la persona que recibe el soborno. Porque el mismo delito, la venta ilegal o el soborno, necesita que dos personas lo cometan, de modo que castigarlo por la razón adicional de que se trataba de una conspiración (lo que requiere un mínimo de dos conspiradores) sería un cómputo doble.

Una regla llamada doctrina Pinkerton, llamada así por el caso que la estableció, responsabiliza a todos los participantes de una conspiración de cualquier delito cometido por cualquier miembro dentro del alcance de ella, siempre que el delito fuera previsible para los demás integrantes. De modo que, si un miembro de una banda de narcotraficantes golpea a un cliente moroso, los otros miembros también son susceptibles de ser considerados culpables de asalto y agresión, a pesar de que no tuvieron nada que ver con la golpiza. Una regla relacionada, la regla del homicidio con delito, culpa a un delincuente de asesinato en primer grado si un homicidio ocurre en el curso de su delito, incluso si el asesinato lo comete otra persona y él no lo autorizó o ni siquiera lo esperaba, como cuando un policía mata accidentalmente a un transeúnte en el curso de su intento por frustrar un delito.

Éstos son ejemplos de castigo colectivo, una práctica que, como señalamos en el capítulo VI, tiene sentido económico cuando alguien que no es el autor material de un acto ilícito puede tener más información que el gobierno, la cual podría utilizar para evitar la acción, si se le motivara. Los conspiradores tienen un incentivo para vigilarse entre sí a fin de evitar meterse en problemas innecesarios; y los autores del robo de un banco tienen un incentivo para evitar estar armados o provocar a los guardias del banco o a la policía.

Estrechamente relacionada con el delito de conspiración se encuentra la instigación y la complicidad. Comparemos dos casos: *1)* un comerciante le vende un vestido llamativo a una mujer de quien sabe que es prostituta; *2)* un comerciante le vende una pistola a un hombre que le dice que está planeando utilizarla en un homicidio. En ambos casos, la responsabilidad penal aumentará los costos esperados del delincuente (principal) sólo ligeramente, porque la prostituta o el homicida incurrirán en escaso costo añadido comprando en tiendas que no conocen la ocupación de ella o las intenciones de él. En el segundo caso, sin embargo, los beneficios de la imposición de una responsabilidad penal al comerciante son mayores porque el delito del infractor principal es mucho más grave, y ésta es la situación en la que la ley impone la responsabilidad.[22]

El entrampamiento se relaciona estrechamente con la tentativa, aunque es una defensa contra un juicio penal, mientras que la tentativa es un delito. Por ejemplo, la policía le solicita a menudo a una persona que cometa un delito. La forma más común de esta táctica es el envío de un agente encubierto o de un informante para que le compre drogas a un narcotraficante, quien es luego "atrapado" y procesado por una venta ilegal. Podría parecer extraño que la ley castigara un acto tan inocuo, porque es obvio que la venta

[22] ¿Por qué no son también coconspiradores todos los que instigan y son cómplices? Porque la conspiración requiere un acuerdo. El caso más claro de un instigador y cómplice que no era también un coconspirador sería el de alguien que apareció por azar en un robo en proceso y, siendo un anarquista, le metió una zancadilla a un guardia de seguridad que estaba tratando de impedir la consumación.

de drogas a un agente encubierto que luego destruye las drogas no causa daño a nadie. Lo único importante, podría parecer, sería recuperar del vendedor el dinero usado en la compra. Pero la justificación es de nuevo la prevención.[23] Éste es un acto inocuo, pero es muy probable que el vendedor realice ventas ilegales si no es arrestado, de modo que lo arrestamos y condenamos ahora porque resulta mucho más barato atraparlo en un delito arreglado que en sus actividades criminales ordinarias. Los beneficios del encarcelamiento son virtualmente iguales, pero los costos de la aprehensión y la condena son mucho menores.

Esta clase de entrampamiento es perfectamente legal. La defensa del entrampamiento opera sólo si la persona entrampada no estaba "predispuesta" a cometer el delito: predispuesta en el sentido de que habría cometido el mismo delito, sólo que bajo circunstancias que habrían dificultado más su aprehensión por la policía, si no hubiese caído en la trampa policiaca. Pero supongamos que, en lugar de simular las oportunidades delictivas normales del sujeto en la mira, la policía le ofrece alicientes tales que lo persuadieran de cometer delitos que jamás cometería en su ambiente ordinario. Por ejemplo, él es un ladronzuelo y la policía le ofrece 100 000 dólares para robar un caballo pura sangre. Los alicientes de la policía que sólo afecten al momento oportuno y no al nivel de la actividad criminal son socialmente productivos; las que inducen un nivel más alto de tal actividad no lo son.[24]

Un método favorito de entrampamiento legal es la "trampa de la casa de seguridad". Una casa de seguridad es un lugar donde se almacenan drogas ilegales. En la trampa de la casa de seguridad, la policía persuade a los narcodelincuentes para que acepten robar una casa de seguridad, pero no existe tal casa; sin embargo, si aceptan y se satisface el requisito de la prueba de predisposición, se les castiga severamente por conspiración. Debido a que la gravedad de la pena aumenta según la cantidad del robo previsto, la policía inventa una cifra grande para el valor del inventario de drogas ilegales de la (falsa) casa de seguridad.

Esta forma de entrampamiento es perversa; es más probable que aumente la venta de drogas ilegales en vez de reducirla. Los ladrones de la casa de seguridad no aumentan la cantidad de drogas en circulación, ya que roban sus drogas en vez de fabricarlas o importarlas. Por tanto, el efecto de la trampa de la casa de seguridad falsa es volver más seguras las casas de seguridad al reducir la probabilidad de que sean robadas. La trampa elimina a algunos ladrones potenciales de casas de seguridad y disuade a otros de aceptar participar en robos a casas de seguridad, ya que los "robos" pueden resultar ser trampas. La mayor seguridad que la trampa de la casa de seguridad falsa

[23] Aunque hay también una justificación relacionada con la disuasión. El delincuente se vuelve incierto si el vendedor, en el ejemplo del texto, es un vendedor de buena fe o un agente encubierto.

[24] Véanse United States *vs.* Hollingsworth, 27 F.3d, 1196 (7th Cir., 1994) (en banc); United States *vs*. Wise, 221 F.3d, 140 (5th Cir., 2000).

brinda a las casas de seguridad reales —una seguridad obtenida sin costo alguno para los operadores de casas de seguridad— reducirá su costo de autoprotección, que es un costo principal del negocio ilegal de drogas. Cuanto más bajos sean los costos de una empresa, menores serán los precios que se cobren a los consumidores, por lo que la demanda de drogas ilegales será mayor al igual que sus ventas y consumo. Los operadores de las casas de seguridad *pagarían* el cumplimiento de la ley para provocar a los ladrones potenciales de las casas de seguridad.

La distinción destacada en esta sección entre la disuasión y la prevención tiene contrapartes en otras áreas del derecho. Por ejemplo, podríamos recurrir al sistema de los cuasidelitos para disuadir a los restauranteros de servir alimentos descompuestos; pero no lo hacemos. En lugar de ello, requerimos que los restaurantes tengan permisos y se sometan a inspecciones. La diferencia se encuentra entre la regulación *ex ante* y la regulación *ex post*. Entre más graves sean las consecuencias si falla la disuasión, y, por ende, más severa sea la sanción óptima *ex ante*, más vigoroso será el argumento económico en favor de la regulación *ex ante*. (Más sobre este punto adelante, en VII.7.)

VII.4. Los perdones

Hasta ahora hemos hecho hincapié en los procedimientos que tratan de volver más costoso el delito para el delincuente. ¿Podrá el economista darle algún sentido al poder establecido del presidente (en el caso de los delitos federales) y de los gobernadores (en el caso de los delitos estatales) para perdonar a delincuentes? Aunque los perdones pueden otorgarse en cualquier momento posterior a la comisión de la infracción —aun antes de que se presente la acusación—, se otorgan más a menudo después de que el infractor ha purgado su sentencia. ¿Por qué desearía alguien un perdón cuando ya ha purgado su sentencia? La respuesta depende en parte del efecto legal de un perdón. Una concepción es que ocurre aquí como si nunca hubiera ocurrido la infracción. La infracción se borra del registro del infractor, sus derechos civiles se restablecen (el derecho de votar, el derecho de tener armas, etc.), la condena no puede emplearse para impugnar su testimonio en un juicio (véase después XXIII.6) y la conducta subyacente a la infracción no puede esgrimirse como una prueba en su contra en un procedimiento civil (expulsarlo de la barra, por ejemplo, si se trata de un abogado). Es como si nunca hubiese sido condenado en primer lugar. La otra concepción, actualmente en auge, es que el único efecto de un perdón es la eliminación de las incapacitaciones estatutarias expresas de una condena, tales como la negación del derecho de votar o la descalificación para obtener contratos gubernamentales.

Sin embargo, éstas son sólo las consecuencias *legales* de un perdón, y las consecuencias no legales podrían ser más importantes para muchos indivi-

duos perdonados. Dado que algunos individuos son perdonados debido a una creencia de que fueron injustamente considerados culpables o sentenciados, un perdón lleva cierta connotación de reivindicación, aunque, en efecto, muchos perdones se otorgan por razones muy diferentes, principalmente políticas, si bien se otorgan a veces en reconocimiento del buen comportamiento del individuo después de la condena. Puesto que no existe un criterio claro para otorgar el perdón, los individuos que reciben perdones por razones no relacionadas con la justicia de su condena o con su buen comportamiento después de la condena salen beneficiados sin esfuerzo de los receptores de perdones otorgados por "méritos".

La disponibilidad de los perdones por buen comportamiento después de la condena opone la reincidencia con una zanahoria para complementar al palo más común (un castigo más severo para los infractores reincidentes que para quienes delinquen por primera vez). Por ello, mientras los perdones reducen el castigo *ex post* y, por ende, el costo esperado del castigo del delito, realmente podrían reducir la tasa de delitos en virtud de su efecto sobre el comportamiento de los delincuentes después de la sentencia. Sucede lo mismo con la conmutación de la pena, lo que implica reducir la duración de la sentencia del demandado antes que reducir cualquiera de las consecuencias colaterales de la condena. La conmutación de la pena también puede aumentar el castigo *ex ante*: los jueces pueden estar menos renuentes a imponer sentencias severas si consideran que el presidente o el gobernador pueden conmutar la sentencia en caso de ser demasiado larga.

VII.5. La intención criminal

La disposición de ánimo del criminal acusado constituye una consideración general del derecho penal. Esto es desconcertante para el economista; podemos leer docenas de libros de economía sin encontrar una sola referencia a la intención. En efecto, el concepto de la intención en el derecho penal desempeña funciones económicas. Una es la de diferenciar las transferencias coercitivas puras de la conducta dotada de un perfil económico que se les parece. Si tomo un paraguas de un restaurante y me lo llevo a casa creyendo que es mío, pero no lo es (y lo devuelvo), no soy un ladrón; pero si sé que el paraguas no es mío y de todos modos lo tomo, sí lo soy. El primero es un caso de negligencia simple, ya que en términos de la Fórmula de Hand B es positiva (tendría que gastar recursos para no tomar el paraguas), y L (la pérdida temporal del propietario del uso de su paraguas) es pequeña; y, en un caso de negligencia simple, es mayor el riesgo de que una sanción penal disuada en exceso. Pero, en el segundo caso, cuando gasto recursos para tomar el paraguas de alguien (quizá fui al restaurante con el único propósito de robar un paraguas), B es negativa y P es grande (véase antes VI.15). El proble-

ma es que los actos externos implicados en estas dos transacciones son los mismos, y la disposición del ánimo con el que se hicieron puede ser la única clave para su diferencia económica.[25]

Pero debemos tener cuidado de distinguir entre la intención y la conciencia. De otro modo podríamos caer en la trampa de pensar que los administradores de un ferrocarril son homicidas porque saben casi con certeza que sus trenes se toparán este año con cierto número de personas en los cruces del ferrocarril. Lo saben, pero no obtienen ningún beneficio de las muertes. Sólo obtienen un beneficio del ahorro de los recursos necesarios para prevenir las muertes; y el beneficio, social y privado, podría superar al costo. La intención relevante para la responsabilidad penal es la intención de alcanzar cierto objeto (prohibido) invirtiendo recursos en su obtención.

Cuando los costos sociales de evitar alguna actividad lícita por miedo al castigo penal erróneo se consideran menores, la ley no insiste en la prueba de la mala intención. Un ejemplo famoso es el de la violación estatutaria. La muchacha podría parecer de 17 años (supongamos que tal es la edad del consentimiento), pero, si es menor, un error razonable no excusará al varón. Otro ejemplo es del homicidio con delito. En estos ejemplos no nos interesa la desistencia de la actividad en el barrio donde se realiza la actividad a la que se dirige la prohibición penal básica (delitos, en el caso de homicidio); en otras palabras, no contamos la evitación de esa actividad como un costo social, de modo que conviene reducir el costo del procesamiento eliminando la cuestión de la intención. Introducimos cierto grado de responsabilidad estricta en el derecho penal cuando, como en el derecho de los cuasidelitos, un cambio en la actividad es un método eficiente para evitar un costo social (véase antes VI.5).

El homicidio premeditado se castiga más severamente que el homicidio cometido en un arranque de ira. La justificación económica es que es probable que el individuo que planea un homicidio planee al mismo tiempo su huida, mientras que es probable que un homicidio cometido al calor de la pasión ocurra en circunstancias en las que el ocultamiento no ha sido considerado ni se puede realizar. Por lo tanto, p en nuestra fórmula $D = L/p$ es menor en el primer caso, lo que implica que el castigo debiera ser más severo. Sin embargo, adviértase que el estado de ánimo requerido es la deliberación intencional antes que el simple conocimiento de los hechos que convierten al apoderamiento en una transferencia coercitiva.[26]

[25] Pero, a menos que el delincuente confiese o haga otras admisiones comprometedoras, su disposición de ánimo debe inferirse también de los actos externos. Cuando el peligro es evidente no escucharemos a alguien que dice que no es consciente de ello, a menos que estemos convencidos de que su racionalidad estaba tan deteriorada que para él la apariencia era completamente diferente de la que sería para nosotros, como cuando un loco mata a su madre pensando que era un mosquito. Por lo tanto, la intención se puede entender como la locución de un lego para una proporción elevada de *PL* a *B*. *Cf.* Antes en VI.15.

[26] Jeffrey S. Parker destaca esta distinción en "The Economics of Mens Rea", 79 *Va. L. Rev.*, 741 (1993).

Otra razón para castigar el delito impulsivo menos severamente que el homicidio deliberado es que el criminal impulsivo está menos sujeto a la disuasión; por lo tanto, el castigo es menos eficaz, de modo que la sociedad debiera comprar menos de él. Empero, aunque es cierto que el delincuente que es suficientemente reflexivo para planear su delito previamente tendrá mayores probabilidades de ponderar el alcance completo de los costos y los beneficios asociados al delito, incluido el costo esperado del castigo, el hecho de que un aumento dado del castigo disuadirá al delincuente impulsivo menos que al deliberado podría aconsejar en efecto un castigo *más severo* para el primero. Supongamos que una sentencia de 20 años es suficiente para disuadir a la virtual totalidad de los homicidas a sueldo, pero el logro de la misma disuasión de los homicidas impulsivos requeriría 30 años. El costo para el gobierno de la sentencia con tiempo adicional puede ser menor al beneficio social de la disuasión adicional generada. Y no olvidemos el efecto preventivo del encarcelamiento. El hecho de que ciertos delincuentes no sean efectivamente propensos a la disuasión aconseja un énfasis mayor en su incapacitación, lo que implica sentencias de prisión largas.

El conflicto entre la disuasión y la prevención como objetivos de la sanción penal es especialmente agudo en relación con la defensa de la locura. Si una persona está loca en el sentido de que no sabe que lo que está haciendo es un delito (mata a un hombre creyendo que es en realidad un hombre lobo) o de que no puede controlarse (escucha voces que toma por mandatos divinos para que mate gente), no se verá disuadida por la amenaza del castigo penal. Por lo tanto, si la disuasión fuese todo el efecto de la sanción penal, tales personas no debieran ser castigadas como delincuentes;[27] los recursos consumidos en su castigo (incluida la desutilidad del castigo para el "criminal" mismo) serían un desperdicio porque no comprarían nada de disuasión. En realidad, ésta es una formulación exagerada: la existencia de una defensa de locura atraerá recursos para probarla y refutarla, y la disuasión se verá disminuida en la medida en que los delincuentes logren engañar fingiendo la locura o en que una disminución (por cualquier razón) del número de las personas castigadas reduzca la señal disuasiva emitida por el castigo. Pero, dejando de lado todo esto, una vez que ocupa la atención la meta preventiva resulta mucho menos claro que debiera haber una defensa de locura, porque aumenta el costo del proceso penal sin reducir en lo más mínimo la necesidad de incapacitar al acusado.

Sin embargo, el efecto de estigma del castigo penal disminuiría si fuese punible una clase completa de individuos no capaces de disuadirse. El estigma

[27] ¿Se aplica esto también si la defensa de la locura sólo requiere que el delito del demandado sea el producto de su locura (la regla de Durham), en el sentido de que si no hubiera sido por su locura no lo habría cometido? No: incluso si el delincuente no hubiese cometido el delito de no ser porque está loco, es posible que el castigo lo disuada bastante, y castigarlo puede ser más barato que curar su locura.

depende de la percepción popular de que el castigo penal se impone a las personas malas, y al menos algunas personas consideran que las personas que matan porque están locas están enfermas, mas no que son malas. Éste es un argumento en favor del empleo del derecho civil, antes que del derecho penal, para incapacitar al individuo penalmente loco si se considera que el estigma es un componente importante del castigo penal. En este análisis, el reconocimiento de la defensa por locura aumenta paradójicamente el efecto disuasivo del castigo penal.

La locura se reconoce raras veces como una defensa en el derecho aplicable a los cuasidelitos, ya que el estigma no es una característica importante del arsenal correctivo de esta rama del derecho; y, si una persona demente posee bienes, su familia u otros cuidadores tendrán un incentivo para evitar que cometa agravios.

Algunos problemas de la intención criminal se esclarecen por el concepto de los costos de la información. Por ejemplo, a menudo no está claro si el comprador de objetos robados sabía que lo eran. La prueba de su responsabilidad penal consiste en saber si, sospechando que los bienes eran robados, conscientemente evitó la adquisición del conocimiento que habría verificado o despejado sus sospechas. Esta prueba impone, en efecto, al comprador una obligación legal, exigible mediante el castigo penal, de investigar la procedencia de los bienes cuando los costos de la investigación son muy bajos.[28] O consideremos la máxima de que la ignorancia de la ley no es una defensa contra la responsabilidad penal. En virtud de que las leyes penales poco claras pueden generar sustanciales costos (por mantenerse alejado) para evitar el delito, estas leyes son generalmente claras, si bien menos porque estén redactadas con claridad que por confinarse a un tipo de conducta que todos reconocen como antisocial. El resultado es un costo en extremo bajo de la adquisición de conocimientos acerca de nuestras obligaciones bajo el derecho penal. Si el costo en cuestión es elevado, el derecho penal requiere de ordinario la prueba de la "intención específica", es decir, el conocimiento no sólo de los hechos que establecen el delito sino también el de que *se trata de* un delito. La evasión del impuesto al ingreso es un ejemplo. Si una persona sabe que debe más impuestos, *B* en la Fórmula de Hand es cero. Si no lo sabe, *B* podría ser sustancial, ubicando al caso en el intervalo en el que los remedios de los cuasidelitos son preferibles.[29]

[28] Los costos pueden ser incluso negativos, como ocurre cuando el que recibe los bienes robados gasta efectivamente recursos para evitar la adquisición del conocimiento culpable; por ejemplo, podría seguir una ruta indirecta a fin de no ver el robo en progreso. Véase United States *vs.* Giovannetti, 919 F.2d, 1223 (7th Cir., 1990).

[29] Véase un análisis más detallado del problema de la ignorancia en el derecho penal en Louis Kaplow, "Optimal Deterrence, Uninformed Individuals, and Acquiring Information About Whether Acts Are Subject to Sanctions", 6 *J. Law Econ. & Organization,* 93 (1990).

VII.6. Los delitos de odio

En la actualidad, los delitos motivados por el odio a un grupo particular tal como el de los negros o los homosexuales tienden a castigarse con mayor severidad que los mismos delitos cometidos por otras motivaciones. Si grupos particulares, tales como los negros del Sur durante la época de Jim Crow, tienden menos que otras víctimas del delito a reportar un delito en su contra o a recibir una protección eficaz de la policía y otras autoridades encargadas de aplicar la ley si lo reportan, entonces serán mayores los beneficios esperados por los delincuentes del hecho de victimar a los miembros del grupo, lo que ameritará un castigo más severo. También se justifica el castigo más severo si el daño psicológico de la víctima de un crimen de odio es mayor cuando la víctima sabe que el criminal está motivado en todo o en parte por el odio al grupo al cual pertenece la víctima antes que por la avaricia o la hostilidad personal, o si tales crímenes son difíciles de evitar por las víctimas potenciales (véase después VII.7). Pero los beneficios del castigo incrementado deben evaluarse junto con los costos para el gobierno de tener que probar un elemento adicional del delito, o sea, la motivación del acusado.

Las leyes de los delitos de odio son criticadas a veces por individualizar a grupos específicos para una protección mayor de las leyes penales. Una persona que asesina a homosexuales porque los odia es más peligrosa que una persona que asesina al hombre que se ha acostado con su mujer, pero no más peligrosa que alguien que asesina prostitutas porque las odia. Otra crítica consiste en que las leyes de los delitos de odio castigan el pensamiento antes que la acción. Comparemos a dos chantajistas de homosexuales. Uno de ellos comete el chantaje puramente por dinero y el otro lo hace en parte movido por su odio a los homosexuales.[30] El chantaje a cargo del primer chantajista no es un crimen de odio; el chantaje a cargo del segundo sí lo es. Castigar al segundo más severamente puede parecer que castiga (hasta un punto medido por el aumento del castigo del segundo chantajista sobre el primero) una *opinión* acerca de la homosexualidad. No obstante, puede ser un delito mucho más dañino para la víctima debido a la motivación del delincuente. Esta inquietud se refleja en la directriz federal para las sentencias que insta a aplicar un castigo mayor para los delitos contra "víctimas vulnerables".[31]

[30] Supongamos que su víctima ignora su motivación, ya que una manifestación de la homofobia del segundo chantajista podría aumentar el daño psicológico de su víctima.

[31] Véase, por ejemplo, United States *vs.* Lallemand, 989 F.2d, 936 (7th Cir., 1993), que confirma el aumento cuando la víctima del chantaje era un homosexual… pero lo confirma no porque al demandado lo motivara el odio hacia los homosexuales (al parecer, su motivación era estrictamente financiera), sino porque se trataba de un homosexual refugiado en lo más recóndito del clóset y, por ello, difícilmente levantaría una queja, lo que redujo el costo del castigo esperado al chantajista.

VII.7. Descuido, negligencia y responsabilidad estricta

Hemos señalado que, en virtud de que las sanciones penales son severas, su imposición a la conducta accidental (y *a fortiori* a la conducta inevitable) genera incentivos para alejarse de actividades legítimas a fin de evitar el riesgo del castigo penal erróneo. Pero, entre las excepciones, la conducta altamente imprudente se equipara a veces, tanto en el derecho penal como en el de los cuasidelitos, a la conducta deliberada. Si X sólo por el placer de hacerlo dispara con un rifle a las ventanas iluminadas de un tren en movimiento y un pasajero del tren es asesinado, X será culpable de asesinato en primer grado. En este caso, tanto *P* como (especialmente) *L* son grandes, y B es negativa porque X ha gastado recursos reales para poner en riesgo a los pasajeros. Desde luego, hay una compensación: la satisfacción que X obtiene del disparo. X no podría haber obtenido la misma satisfacción apuntando a un blanco en el campo de tiro, porque la satisfacción que obtuvo de ello en el juicio dependía, sin duda, en cierta medida de poner en riesgo la vida de las personas. Eso hace que sea una variante del caso de malicia (utilidades negativas interdependientes) discutido antes en VI.15, y, por las razones que se explicaron allí, este tipo de satisfacción no debe contar en el cálculo del bienestar social.

Consideremos el asesinato en la creencia honesta pero poco razonable de que ello es necesario en defensa propia. Éste es un asesinato deliberado, así que *P* y *L* son grandes. *B* es grande también, no obstante; por definición, el homicida teme por su propia vida. Sin embargo, la brecha entre *PL* y *B* bien puede ser sustancial, lo que, aunado al hecho de que *L* es grande, justifica el castigo penal de una conducta que en un sentido importante es accidental. En el ejemplo, el delito sería homicidio involuntario, no homicidio simple; la brecha existente entre *PL* y *B* es menor que en el caso del homicidio por descuido que se castiga como un homicidio de segundo grado, por lo que el castigo debiera ser menos severo.

Incluso la negligencia simple se considera algunas veces como delito, e incluso hay delitos de responsabilidad estricta, como la venta de alimentos adulterados, aun cuando el vendedor tomó todas las precauciones justificadas por el costo. La explicación involucra la diferenciación entre las sanciones *ex ante* y *ex post*,[32] de modo que nos encamina hacia el análisis de la regulación de la seguridad en el capítulo XIV. En el caso de la conducta que pone en peligro la vida —un elemento de la mayoría de los delitos de responsabilidad estricta en los que la responsabilidad se basa en la negligencia simple o es estricta— tratar de compensar a la víctima en un juicio de cuasidelito podría resultar un acto inútil, porque el cuasidelincuente carece de dinero para

[32] Véase Steven Shavell, "Liability for Harm *versus* Regulation of Safety", 13 *J. Legal Stud.,* 357 (1984).

pagar una condena tan grande. Dado que el sistema de los cuasidelitos no puede desempeñar un papel disuasivo ni compensar a la víctima, la disuasión se vuelve responsabilidad del gobierno. Éste es otro ejemplo de la profunda relación entre el derecho penal y las limitaciones de solvencia.

En Commonwealth *vs*. Mixer,[33] el conductor de una compañía de camiones fue castigado por violar una ley que prohibía transportar licor en ciertas localidades a las compañías de transportes y a sus empleados. El tribunal explicó que la legislatura había agotado todas las posibilidades para evitar el transporte ilegal de licor, había dudado de que hubiera muchos casos en los que la compañía o su empleado no se dieran cuenta de que aquélla o éste transportaban licor ilegalmente y había esperado que, al imponer la responsabilidad penal estricta, se indujera a los transportistas comunes y a sus empleados a ser más cuidadosos. El análisis del tribunal capturó con tino las tres razones principales para imponer, en ocasiones, la responsabilidad penal estricta: la dificultad de demostrar el conocimiento, el escepticismo ante la versión de que los demandados realmente ignoraban los hechos críticos y el incentivo para tener el cuidado extra que crea la responsabilidad estricta.

Otro ejemplo: la ley federal establece lo siguiente: "En caso de muerte o lesiones corporales graves derivadas del uso de [drogas ilegales de parte del demandado], [éste] será condenado a prisión por un plazo no menor a veinte años o mayor que la cadena perpetua".[34] Esta responsabilidad estricta crea incentivos para que el traficante de drogas advierta a su cliente sobre la potencia del lote específico de drogas que se están vendiendo y se niegue a suministrar drogas a las personas especialmente vulnerables. Y, en este caso, la responsabilidad estricta no viola el principio de la disuasión marginal porque no da un incentivo al vendedor para cometer un delito más grave, como en el caso en el que el robo se castiga con tanta severidad como el asesinato.

¿Pero exactamente qué significado debe atribuirse a "derivadas de"?[35] Supongamos que un demandado vende una droga ilegal a una persona que, al no querer ser vista ingiriéndola, lo hace en su baño y mientras está allí el techo del baño se derrumba y la mata. Si no hubiera comprado la droga no habría perdido la vida. Pero la venta de la droga no aumentó el riesgo que representaba el techo inseguro. Castigar a los vendedores de drogas no detiene los accidentes en los edificios.

VII.8. La defensa de la necesidad (compulsión)

El notorio caso de Regina *vs.* Dudley y Stephens[36] implicó un juicio de homicidio de dos hombres que, *in extremis* en un bote salvavidas, mataron y se

[33] 94 N.E. 249 (Mass., 1910).
[34] 18 U.S.C. 841(b)(I)(G).
[35] Véase United States *vs*. Hatfield, 591 F.3d, 945 (7th Cir., 2010).
[36] 14 Q. B. D., 273 (1884).

comieron a uno de los miembros del grupo, un grumete. Se hizo valer una defensa de necesidad o compulsión, pero fue rechazada, aunque después de haber sido sentenciados a muerte se indultó a los demandados enseguida. En el derecho moderno esta defensa, que sigue siendo considerada con renuencia excepto cuando asume la forma de la defensa personal, de ordinario triunfará si hay una disparidad muy grande entre el costo del delito para la víctima y la ganancia para el victimario. Adviértase también que, al revés de lo que ocurre en el caso de la locura —un tipo de defensa fundamentalmente diferente—, no se serviría a ninguna meta incapacitante rechazando una defensa de necesidad; en este caso no *deseamos* la incapacidad.

Pero cambiemos ligeramente el ejemplo: me estoy muriendo de hambre e imploro una pieza de pan a un rico goloso que me la niega. Si sigo adelante y arranco el pan de su mano, seré culpable de robo y no podré interponer una defensa de necesidad. La justificación económica de este resultado cruel es que, puesto que los costos de transacción son bajos, mi imposibilidad para negociar una compra exitosa del pan revela que el pan vale realmente más para el goloso. Los costos de transacción eran prohibitivos en nuestro ejemplo anterior del robo de alimentos de una cabaña en el bosque, de modo que ése es un mejor argumento para la defensa.

Volvamos ahora a Dudley y Stephens. Había pruebas de que el grumete estaba a punto de morir de todos modos, y de que darle muerte y devorarlo salvó las vidas de los demandados y de otro hombre en el bote (quien no fue acusado porque no había participado en la matanza). Sin embargo, a menos que la víctima supiera que se iba a morir de cualquier forma, es probable que no hubiera vendido su vida a los otros a ningún precio. Por lo tanto, este caso parece similar al del mendigo que se muere de hambre. Pero algo debe de estar mal. Aunque los costos de transacción no fuesen elevados en el sentido habitual, la mayoría de nosotros pensaríamos que en cierto momento el sacrificio de una persona para que los demás vivan aumenta el bienestar social. Si pudiera demostrarse que antes de iniciar el viaje los miembros de la tripulación habían convenido en el sacrificio del más débil si ello se hiciera necesario para salvar a los demás, habría un argumento económico a favor de que se permitiera la defensa de la necesidad.[37]

¿Sería un volado un método mejor o peor para seleccionar a la víctima, desde un punto de vista económico? (Adviértase la analogía, por extraña que pueda parecer, con la regla marítima del promedio general, discutida antes, en VI.4, nota 17.) Esto proporcionaría más protección del seguro para los débiles y también es un criterio un poco menos manipulable. Sin embargo, puede derivar en una pérdida: el sacrificio de una persona fuerte por una persona débil que está muy lejos de salvarse, dando como resultado dos muer-

[37] Véase A. W. Brian Simpson, *Cannibalism and the Common Law,* 140, 145, 233-234 (1984); United States *vs.* Holmes, 26 Fed. Cas. 360 (Cir. Ct. Pa. 1842).

tos en una situación en la que un método de selección diferente habría dado como resultado uno solo.

VII.9. Precauciones de las víctimas; acerca de la provocación y el control de armas

Al discutir el tratamiento de las transferencias coercitivas puras en el derecho aplicable a los cuasidelitos, vimos que no es eficiente el reconocimiento de una defensa de la negligencia contribuyente o comparativa, porque por definición el victimario (quien ejerce la coerción) puede evitar el daño a un costo menor que el de la víctima (véase antes VI.15). En efecto, el costo de evitar el acto, aunque positivo para la víctima, es negativo para el victimario. ¿Debiera aplicarse el mismo principio en el derecho penal, por lo menos cuando ese derecho está castigando las transferencias coercitivas puras? Quizá no.[38] En virtud de que las sanciones penales son más costosas que las sanciones de los cuasidelitos, no es posible omitir el costo del empleo del derecho penal al decidir si deberá imponerse o no una obligación a las víctimas. Aunque el delincuente puede evitar el perjuicio de la víctima a un costo menor que la propia víctima, la imposición de la responsabilidad a la víctima podría minimizar los costos sociales agregados. Cuesta algo, aunque muy poco, cerrar con llave nuestro automóvil… quizá menos que lo que le costaría al sistema de la justicia penal llevar al ladrón ante la justicia.

Aunque el derecho penal no reconoce una defensa plena de negligencia de la víctima, sí reconoce una defensa parcial (equivalente a la defensa de la negligencia comparativa en el derecho aplicable a los cuasidelitos) bajo el nombre de provocación. Quien provoca un delito contra sí mismo debiera saber que el delincuente no será castigado tan severamente, lo que implica una menor desistencia de los delitos provocados, lo que a su vez otorga a la víctima un incentivo mayor para evitar la provocación. Y recordemos que las directrices para las sentencias federales proveen una sentencia más severa cuando la infracción se comete contra "víctimas vulnerables", una clase que no incluye sólo a personas que probablemente resultarán lesionadas más gravemente por el acto criminal sino también a personas que carecen de la habilidad media para evitar convertirse en una víctima del delito.

Las precauciones de la víctima, como las alarmas contra robos o el hecho de cerrar con llave las puertas de nuestro automóvil, son criticadas a veces por trasladar el delito en lugar de reducirlo. Y es cierto que la precaución impone un costo externo al volver a las personas que no toman la precaución en objetivos más atractivos del delito. Sin embargo, la validez de esta crítica depende de que la precaución sea o no evidente para el delincuente antes de

[38] Véase Alon Harel, "Efficiency and Fairness in Criminal Law: The Case for a Criminal Law Principle of Comparative Fault", 82 *Cal. L. Rev.,* 1181 (1994).

que cometa el delito. Si la respuesta es negativa, el efecto de la precaución aumenta de manera inequívoca el costo esperado del delito. Lojack, un radiotransmisor oculto instalado en un automóvil, permite que la policía determine su ubicación. Lojack reduce drásticamente la incidencia del robo de automóviles, pero se ha estimado que se externalizan 90% los beneficios de la disminución de los robos, es decir, se transfieren a las personas que no tienen Lojack.[39] Los potenciales ladrones de automóviles, ignorando si un automóvil está equipado con Lojack, tienen menores probabilidades de robárselo aunque no esté equipado así. Dado que los beneficios sociales de Lojack superan ampliamente sus beneficios privados, ¿debiera éste subsidiarse o hacerse obligatorio? No necesariamente; si suficientes personas tienen Lojack para disuadir a la mayoría de los ladrones de automóviles debido a la incertidumbre acerca de si un auto específico tiene el dispositivo o no, los beneficios incrementales de subsidiar u obligar a todos los demás conductores o a la mayoría de ellos para instalar el dispositivo pueden ser menores al costo adicional. Pero, si se supiera que muy pocos coches estaban equipados con Lojack, que la mayoría de los propietarios de automóviles eran oportunistas, el efecto disuasivo del dispositivo se reduciría; un ladrón sabría que existe una gran probabilidad de que el auto que quiere robar no cuente con este dispositivo.

La más controversial precaución de la víctima contra el delito es la posesión de un arma de fuego. Las leyes del control de armas de fuego tienen dos efectos opuestos: *1)* vuelven más costosa para los delincuentes la obtención de armas de fuego, elevando así el costo esperado del delito, pero *2)* al volver más costosa para las víctimas de los delitos la obtención de armas de fuego, las leyes reducen el costo esperado de la comisión de un delito. Empíricamente, ningún efecto parece ser de escala mayor.[40] Aunque ahora la Corte Suprema

[39] Ian Ayres y Steven D. Levitt, "Measuring Positive Externalities from Unobservable Victim Precaution: An Empirical Analysis of Lojack", 113 *Q. J. Econ.*, 43 (1998).

[40] "Con base en los resultados de las evaluaciones de la legislación nacional, las comparaciones entre países y los estudios de índices, la evidencia es insuficiente para determinar si el grado o intensidad de la regulación de armas de fuego se asocia con la disminución (o aumento) de la violencia." Robert A. Hahn *et al.*, "Firearms Laws and the Reduction of Violence: A Systematic Review", 28 *Am. J. Preventive Med.*, 40, 59 (2005); *cf.* John J. Donohue, "The Impact of Concealed-Carry Laws", en *Evaluating Gun Policy Effects on Crime and Violence*, 287, 314-321 (Jens Ludwig y Philip J. Cook [comps.], 2003). "Ya sea que el efecto neto de la relajación de las leyes de portación oculta sea el aumento o la reducción de la carga de delitos, existe una buena razón para creer que la captura no es mayor […] [E]l cambio en la portación de armas parece concentrarse en las zonas rurales y suburbanas en las que los índices delictivos ya son relativamente bajos, entre las personas cuyo riesgo de victimización es relativamente bajo: hombres blancos de clase media de mediana edad. Los datos disponibles respecto a los titulares de permisos también implican que su riesgo de hacer un mal uso de las armas de fuego es bastante bajo, de acuerdo con las tasas de arresto relativamente bajas observadas hasta la fecha para los titulares de permisos. Por lo tanto, con base en los datos empíricos disponibles, si los tribunales invalidan las leyes que prohíben la portación de armas de fuego fuera del hogar se espera un impacto relativamente pequeño sobre la seguridad pública, suponiendo que se permita que siga vigente algún tipo

ha interpretado la Segunda Enmienda a la Constitución para crear el derecho a poseer un arma para autodefensa, ha dejado en claro que sólo el respetuoso de la ley tiene ese derecho. Pero mientras más gente respetuosa de la ley tenga armas, más criminales las tendrán, tanto para compensar las capacidades defensivas mejoradas de las víctimas de los delitos como porque, mientras más armas haya en circulación, mayor será el acceso que tengan a ellas los criminales a través del robo.

Los partidarios de las leyes de control de armas señalan que la policía apoya esas leyes, pero esto no es evidencia convincente de que esas leyes reduzcan el crimen. Los policías están más seguros si los criminales no tienen armas; si los respetuosos de la ley carecen de armas de fuego, eso puede aumentar la tasa de los delitos, pero no pondrá en peligro a la policía a menos que dé lugar a que más criminales tengan armas.

Cuando las precauciones de la víctima son un medio eficaz para la prevención del delito, la tasa de los delitos puede ser tercamente resistente a los esfuerzos gubernamentales para reducirla.[41] Supongamos que la tasa de la delincuencia es elevada, a pesar de que las víctimas potenciales están tomando muchas precauciones, porque lo elevado de esa tasa les da un incentivo para hacerlo. Si el gobierno trata de disminuir la tasa de la delincuencia dotando de recursos al sistema de la justicia penal, al principio lo logrará: dicha tasa bajará. Pero, a medida que la tasa baje, las víctimas potenciales del delito tomarán menores precauciones porque el beneficio esperado de éstas será menor conforme sea más baja la tasa de delincuencia. En consecuencia, ésta aumentará. Es posible que los gastos públicos en la prevención del delito hayan remplazado simplemente a los gastos privados.

VII.10. La economía del crimen organizado

Los delincuentes se organizan a veces en firmas ilegales que pueden estar conectadas en cárteles nacionales o aun internacionales, operando en campos criminales tales como la usura, la prostitución, los juegos de azar y las drogas, pero también en campos legítimos, y empleando amenazas de violencia y la corrupción de la policía como principales métodos de negocios. Desde una perspectiva económica, algunas de las presuntas características del crimen organizado son realistas, y otras no lo son. Las actividades aso-

de sistema de permisos para la portación pública." Philip J. Cook, Jens Ludwig y Adam M. Samaha, "Gun Control after Heller: Threats and Sideshows from a Social Welfare Perspective", 56 *UCLA L. Rev.*, 1041, 1082 (2009); véase también H. Sterling Burnett, "Texas Concealed Handgun Carriers"; "Law-Abiding Public Benefactors", www.ncpa.org/pdfs/ba324.pdf. Véase también Levitt, "Understanding Why Crime Fell in the 1990s", VII.2, n. 4, en 173-175.

[41] Véase Thomas J. Philipson y Richard A. Posner, "The Economic Epidemiology of Crime", 39 *J. Law & Econ.,* 405 (1996).

ciadas a éste implican primordialmente relaciones voluntarias de comprador-vendedor antes que intercambios forzados. Dado que tales relaciones suponen de ordinario cierto grado de organización y especialización, no nos sorprende encontrar "empresas" antes que meros individuos aislados ocupados en estas actividades. El carácter relativamente conspicuo de una organización, en comparación con un individuo, parecería volverla muy vulnerable. Pero, en virtud de que las actividades del crimen organizado involucran a víctimas voluntarias, la organización corre un peligro de aprehensión relativamente escaso por el método habitual de queja de la víctima. (De aquí la importancia de los agentes encubiertos, las intervenciones telefónicas, las operaciones policiacas a modo de trampa y otros métodos proactivos del cumplimiento forzoso del derecho penal.) La corrupción de la policía se facilita por la naturaleza continua de la actividad criminal; una relación continua da lugar a métodos informales pero eficaces de cumplimiento contractual que pueden sustituir a los métodos legales que no están disponibles en una actividad ilegal, con lo que es más probable que se cumpla con el "contrato" de cohecho.

No es sorprendente que el crimen organizado deba emplear a veces la violencia (y más a menudo la amenaza de la violencia), ya que está prohibido el cumplimiento de sus contratos por medios legales. Puesto que la mayor parte del crimen organizado no tiene víctimas, la violencia se dirige principalmente a los rivales, a miembros desleales de la organización y al cliente moroso ocasional. Sin embargo, las amenazas de violencia son fundamentales para las redes de extorsión (aparte de aquellas en las que el extorsionador es un funcionario de gobierno amenazando con retener un beneficio legal de la víctima o imponerle un costo legal). Dado que las amenazas de violencia son más baratas que la violencia misma, la violencia efectivamente aplicada por los extorsionadores representa un rompimiento en el "mercado" de la extorsión, de la misma manera en que una huelga representa un rompimiento del mercado de relaciones laborales y un juicio es un rompimiento del mercado de resolución de disputas. Así como la incertidumbre acerca del resultado del litigio es una causa importante de los juicios (véase después XXII.5), igualmente la incertidumbre acerca de la riqueza de la víctima de extorsión podría causar violencia si el extorsionador piensa que la víctima se resiste a una demanda por terquedad antes que por la falta de recursos.[42]

Una vez dicho esto, es probable que el crimen organizado emplee la violencia con menor frecuencia que el crimen no organizado. La violencia atemoriza al público y, por lo tanto, incita a mayores esfuerzos de la policía para prevenir la actividad que la genera. Esta consecuencia de la violencia es

[42] Véase Andrew R. Dick, "When Does Organized Crime Pay? A Transaction Cost Analysis", 15 *Int'l Rev. Law & Econ.*, 25 (1995).

externa al delincuente individual pero no a una gran organización criminal. Por lo tanto, la organización tiene un incentivo para frenar las inclinaciones violentas de sus miembros.

La inexistencia de sanciones legales por el incumplimiento de contratos ilegales implica también que el crimen organizado depende de la confianza más que las partes de los contratos legales. Por lo tanto, es de esperarse que las relaciones familiares sean más comunes en las organizaciones criminales que en las no criminales. Más sutilmente, en virtud de que la confianza se fortalece por las transacciones reiteradas, es de esperarse que el crimen organizado desempeñe un papel mayor en las actividades donde el personal tiene una tasa de rotación más baja. Esto podría explicar el hecho de que las organizaciones criminales involucradas en la prostitución operen a menudo burdeles, pero raramente empleen prostitutas callejeras. La rotación de las prostitutas callejeras es mucho más rápida que la de las prostitutas empleadas en burdeles.[43] Es más probable que aquéllas sean arrestadas, y creen así un registro de antecedentes penales que con el tiempo pueda llegar a ocasionar un castigo severo; y el trabajo es más riesgoso, lo que provoca que se salga del mercado con mayor frecuencia.

Tampoco es sorprendente que las organizaciones criminales traten de entrar a negocios legítimos; éstos proveen atractivas oportunidades de inversión para personas que tengan dinero para invertir y habilidades empresariales. ¿Debiera alentarse o desalentarse tal entrada? Por una parte, un método para la reducción de la incidencia del crimen organizado consiste en aumentar el rendimiento esperado de las actividades legítimas alternativas. Por otra parte, en la medida en que las ganancias obtenidas en el crimen organizado pueden invertirse con seguridad en actividades legítimas para obtener ganancias adicionales, el rendimiento esperado del crimen organizado es mayor de lo que sería de otro modo.

Las características menos plausibles en esta descripción superpuesta del crimen organizado son la presunta escala nacional y aun internacional de las operaciones y las presuntas ganancias monopólicas. Las operaciones a gran escala en los campos habituales del crimen organizado encontrarían deseconomías sustanciales por dos razones. Primera, éstos son en su mayor parte campos de la distribución al mayoreo y el menudeo, y la distribución se encuentra muy descentralizada, lo que indica la presencia de las deseconomías de escala. Segunda, la forma encubierta que se ven forzadas a adoptar las operaciones de las empresas criminales les impide establecer la refinada maquinaria de control —asociada a una comunicación voluminosa— que es necesaria para la operación eficiente de empresas grandes en el sector empresarial legítimo.

[43] Véanse Kai A. Konrad y Stergios Skaperdas, "Extortion", 65 *Economica,* 461 (1998); John J. Donohue III y Steven D. Levitt, "Guns, Violence, and the Efficiency of Illegal Markets", 88 *Am. Econ. Rev*. *Papers & Proceedings*, 463 (mayo de 1998).

Sin embargo, un área del crimen organizado donde el monopolio es importante es la de la red de protección, que es una forma de la extorsión. A fin de persuadir a la víctima de que pague la protección, la pandilla debe ser capaz de garantizarle que nadie más lo golpeará o destruirá su propiedad; en otras palabras, la pandilla debe tener un monopolio local de la violencia extralegal.

¿Debiera la política pública favorecer o combatir a la monopolización de los mercados criminales? Cuando la actividad criminal consiste en la venta de bienes o servicios ilegales como las drogas o la prostitución, el efecto de la monopolización es, al elevar el precio de los bienes o servicios, reducir la cantidad consumida (véase después IX.3). Casi nadie cree que las actividades de cumplimiento de la ley puedan eliminar efectivamente estos mercados ilegales; lo que hace la aplicación forzosa de la ley es aumentar el precio y, por ende, reducir el consumo de lo que se vende en estos mercados. Por lo tanto, si los mercados se hacen más competitivos se obstruye el cumplimiento de la ley.

Se ha sugerido que las sanciones legales para los negocios típicos del crimen organizado, tales como la prostitución y las apuestas, tienen el efecto de crear un "arancel" que permite, a quienes estén dispuestos a asumir el riesgo del castigo penal, obtener ganancias monopólicas entrando a tales negocios.[44] Pero, dado que el costo esperado del castigo es un costo del negocio ilegal y debe ser cubierto como otros costos, un precio que incluya este costo no es un precio de monopolio sino un precio competitivo, aunque un precio mayor del que prevalecería si la actividad fuera legal y no implicara costos de castigo.[45] El meollo de la verdad en el argumento del "arancel del delito" es que los precios más altos creados por las sanciones legales pueden atraer la entrada de empresarios del delito que prefieren el riesgo o de personas que tienen habilidades especiales para evitar la aprehensión; para cualquiera de los grupos, los costos esperados del castigo pueden ser bajos y, por tanto, los ingresos altos. Veremos más sobre esto en la siguiente sección.

VII.11. La "guerra contra las drogas"

Se invierten recursos enormes del sistema de cumplimiento de la ley para luchar contra el tráfico de drogas ilegales, tales como la cocaína, la heroína, la mariguana, el LSD, el éxtasis y las anfetaminas, y hay castigos severos —incluso salvajes— contra ese tráfico. Se ha calculado que la legalización de estas drogas (para lo cual sería necesario que se derogaran tanto las prohibiciones federales como las estatales) reduciría los gastos gubernamentales en

[44] Herbert L. Packer, "The Limits of the Criminal Sanction", 277-282 (1968).
[45] Véase antes en V.4, la discusión del mercado negro de bebés.

41 300 millones de dólares al año, con alrededor de dos tercios de los ahorros correspondientes a los gobiernos local y estatal. Los ahorros implicarían reducciones en los gastos en policía, fiscales, juzgados, prisiones preventivas y penales.[46]

Además, un gran costo que la cifra de 41 300 millones de dólares no toma en cuenta son las numerosas muertes que se atribuyen directamente a las actividades del cumplimiento de la ley dirigidas contra el tráfico de drogas. Como vimos en la sección anterior, ya que a las empresas delictivas se les niega la asistencia del derecho contractual y de los cuasidelitos, éstas recurren a la violencia y a la amenaza de violencia para hacer cumplir sus contratos con sus clientes y regular la competencia entre sí. Cuando la policía arresta al líder de un grupo de narcotraficantes, lo que debilita a la pandilla, los grupos rivales pueden invadir su territorio, desencadenando una guerra entre pandillas, y tales guerras se libran con armas y balas de verdad, y entre las víctimas se encuentran transeúntes desafortunados. Asimismo, debido a que las drogas ilegales no están sujetas a regulación por parte de la Administración de Alimentos y Medicamentos (FDA), son muchos los problemas de calidad, a veces con consecuencias mortales.

En vista de los altos costos de la "guerra contra las drogas", ¿debiera castigarse en absoluto la venta de drogas "peligrosas"? Hay tres argumentos a favor del castigo: *1)* las drogas son destructivas para los usuarios; *2)* sin embargo, muchos usuarios lo ignoran (algunos de ellos son niños), y *3)* el consumo de drogas tiene efectos contra terceros: causa accidentes y delitos. El primer argumento no se consideraría en la economía si todos los adictos fueran adultos capaces, pero algunos adictos son niños y algunos adultos adictos adquirieron la adicción cuando eran niños. El segundo argumento tiene una forma económica, porque ¿quién tiene un incentivo para educar a los individuos acerca de los peligros de las drogas? El contribuyente paga por el cumplimiento del derecho penal y, presumiblemente, pagaría por la educación acerca del peligro de las drogas si ése fuese un método más eficaz para reducir su consumo. Como veremos, es posible que no lo sea. El tercer argumento es excelente por lo que toca a los accidentes, porque las drogas han estado implicadas en algunos accidentes de terceros, pero es un argumento débil por lo que toca a los delitos. Las razones principales por las que el consumo de drogas conduce al delito son, primero, que el castigo penal eleva el precio de las drogas por encima del nivel que los adictos pueden pagar sin cometer delitos para aumentar su ingreso, y, segundo, que los narcotraficantes se ven obligados a usar la violencia para hacer cumplir sus contratos porque se les niega el empleo de los remedios legales; de aquí surgen las "guerras territoriales".

[46] Jeffrey A. Miron y Katherine Waldock, *The Budgetary Impact of Ending Drug Prohibition* (2010), www.cato.org/pubs/wtpapers/DrugProhibitionWP.pdf

Los argumentos económicos a favor de la penalización del narcotráfico no son muy impresionantes.[47] Son tan insignificantes, en lo que respecta a la mariguana, que existe un movimiento a nivel nacional para legalizarla incluso para uso recreativo y no sólo para uso medicinal. Podrían emplearse argumentos similares a ésos a favor de mantener las leyes penales contra la mariguana para proscribir las bebidas alcohólicas (cuyos efectos contra terceros, por la vía de los accidentes, son mayores que los de las drogas ilegales, aunque tal vez se deba principalmente a que hay mucho más alcoholismo que consumo de drogas) y aun los cigarros (es mucho menos probable que los cigarros se impliquen en accidentes, con excepción de los incendios provocados por fumar en la cama, pero son más autodestructivos, al menos físicamente —y físicamente adictivos—, que la mayoría de las drogas ilegales). Pero raras veces se emplean tales argumentos. Es posible que esto se deba sólo al hecho de que hay tantos consumidores de alcohol y de tabaco que tienen la fuerza política suficiente para obstruir los esfuerzos en favor de la prohibición de sus hábitos; pero, precisamente porque el alcohol y el tabaco son sustitutos cercanos y legales de las drogas ilegales, es posible que ni siquiera una victoria completa en el combate a las drogas condujera a una disminución significativa de los problemas que lo motivaron.

Comoquiera que sea, tomando como dada la meta de la extirpación de las drogas actualmente ilegales, ¿cuál es el mejor método para alcanzarla? Algunos arguyen que, paradójicamente, tal método es el de la legalización de las drogas, porque la demanda de drogas no es elástica y, por lo tanto, el consumo de drogas no aumentaría (mucho), y porque la legalización eliminaría las ganancias monopólicas en el negocio de las drogas y, por ende, reduciría el incentivo de los vendedores de drogas para abrir paso a sus mercancías, de modo que el consumo de drogas podría en efecto disminuir, aunque el precio estuviese bajando. Ambos argumentos son muy dudosos. La demanda de drogas adictivas a largo plazo parece ser elástica antes que no elástica.[48] La razón es que el adicto racional en potencia sabe que se encuentra (o se encontrará) "enganchado", de modo que toda disminución permanente del precio (como ocurriría con la legalización de las drogas) disminuirá no sólo el costo actual del consumo sino también los costos futuros. En cuanto a las supuestas "ganancias monopólicas" en el narcotráfico, como vimos en la sección precedente, son sólo una compensación por asumir los riesgos del castigo y los otros costos anormales de un negocio ilegal; no son ganancias en absoluto.

[47] Véase Jeffrey A. Miron y Jeffrey Zwiebel, "The Economic Case against Drug Prohibition", *J. Econ. Perspectives*, otoño de 1995, 175.

[48] Gary S. Becker y Kevin M. Murphy, "A Theory of Rational Addiction", 96 *Journal of Political Economy,* 675 (1988); y véase también Becker, Grossman y Murphy, "Rational Addiction and the Effect of Price on Consumption", 81 *Am. Econ. Rev. Papers & Proceedings*, 237 (mayo de 1991).

Un mejor argumento a favor de la legalización es que todo lo que la política pública puede esperar lograr de manera realista es un aumento en el precio de las drogas ilegales, y esto podría conseguirse de forma más barata mediante la legalización y el gravamen que castigando su venta.[49] Un impuesto al consumo no sólo reduce la demanda, sino que genera ingresos para el Estado, mientras que el gasto en el cumplimiento de la ley diseñada para reducir la demanda no genera ingreso alguno que lo compense,[50] sino que provoca gastos adicionales por parte de los traficantes de drogas para contrarrestar los gastos del gobierno. Adviértase que, en la medida en que la demanda sea no elástica en vez de elástica (contrario a la sugerencia anterior), los gastos en el cumplimiento elevarán el precio y no frenarán el consumo proporcionalmente. Cuando la demanda es no elástica, un aumento en el precio eleva los gastos totales. La demanda no elástica significa que un cambio en el precio conduce a un cambio menos que proporcional en la cantidad demandada. Por ejemplo, si un aumento de 2% en el precio genera una reducción de tan sólo 1% en la cantidad demandada, el ingreso total se elevará.

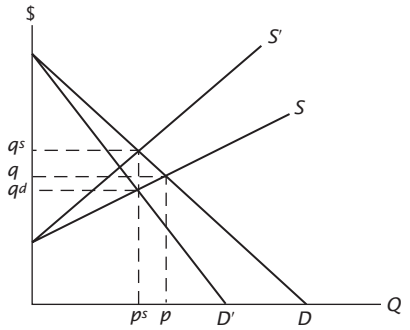
El impuesto puede ser mejor, porque se puede hacer muy severo a un bajo costo. Generará más ingresos para el gobierno entre menos elástica sea la demanda del artículo gravado. Cuanto más elástica sea la demanda, serán menores los ingresos para el gobierno, pero habrá también una mayor sustitución de otros productos.

Por otra parte, a juzgar por los precios actuales de las drogas ilegales, el castigo severo de los traficantes de drogas no ha tenido mucho éxito en el aumento de estos precios. Una de las razones es que, dado que la venta de drogas ilegales es un delito "sin víctimas", en el sentido de que los compradores no tienen incentivos para informar a la policía sobre la venta, la detección de este tipo de delitos y la aprehensión de los delincuentes son tareas costosas, financiadas de manera insuficiente debido a las otras demandas de gasto público. Otra razón es la alta elasticidad de la oferta de traficantes de drogas. Uno es arrestado y otro toma su lugar rápidamente. Las personas a quienes les atrae el tráfico de drogas suelen no tener (o creen que no tienen) oportunidades comparables para ganar un salario digno legalmente. Dada la elasticidad de la oferta, el efecto principal de la detención y el encarcelamiento generalizados de los traficantes de drogas es su rápido remplazo al nivel salarial existente.

Otra posibilidad consistiría en desplazar el énfasis en el cumplimiento de la ley del productor al distribuidor detallista. Supongamos que el costo

[49] Véase Gary S. Becker, Kevin M. Murphy y Michael Grossman, "The Market for Illegal Goods: The Case of Drugs", 114 *J. Pol. Econ.*, 38 (2006).

[50] Los gastos también son muy elevados en comparación con el efecto sobre el precio, que es modesto. Véase Ilyana Kuziemko y Steven D. Levitt, "An Empirical Analysis of Imprisoning Drug Offenders", 88 *J. Pub. Econ.*, 2043 (2004).



GRÁFICA VII.1

de producción de una droga ilegal es 5% del precio de la droga en la calle, pero que el costo de la distribución al menudeo es 50% de ese precio. Si entonces los representantes de la ley que persiguen al productor logran aumentar el costo de producción 20%, el precio en la calle aumentará sólo 1% (20 de 5%). Pero si, al concentrar sus esfuerzos en los distribuidores, los representantes de la ley logran aumentar 20% el costo de los distribuidores, esto se traducirá en un aumento de 10% en el precio de la calle (20 de 50%). Con todo, si hay más distribuidores al menudeo que productores, el costo de su aprehensión podría ser mayor. Por otra parte, uno de los efectos del cumplimiento riguroso de las leyes antidrogas en contra de los distribuidores minoristas puede ser el de inducirlos a reducir el precio de modo que las ventas sean más rápidas (porque mientras más barato sea el producto menos habrá que negociar para su venta) y, por lo tanto, la probabilidad de arresto sea menor.[51] Sin embargo, una reducción en el precio de las drogas, al estimular la demanda, debilita el objetivo del cumplimiento de la ley antidroga.

También se ha sugerido un cambio del énfasis que vaya del castigo a la educación.[52] La gráfica VII.1 compara los dos métodos. El castigo desplaza la curva de oferta hacia la izquierda (de S a S'), lo que conduce a la producción y la venta de una cantidad de drogas menor (q^s) a un precio mayor (p^s). La educación desplaza la curva de demanda hacia la izquierda (de D a D'). De nuevo se produce y se vende una cantidad menor (también q^s), pero se vende a un precio menor (p^d); por lo tanto, disminuye un costo importante del combate a las drogas: la inducción de los adictos a cometer delitos para sostener un hábito caro.

[51] Beth A. Freeborn, "Arrest Avoidance: Law Enforcement and the Price of Cocaine", 52 *J. Law and Econ.*, 19 (2009).

[52] Stephen J. Schulhofer, "Solving the Drug Enforcement Dilemma: Lessons from Economics", 1994 *U. Chi. Legal Forum,* 207.

¿Pero cuán eficaz sería probablemente la educación acerca de los peligros de las drogas? El precio total de las drogas para el consumidor tiene dos componentes: el precio nominal, o sea, el precio que cobra el distribuidor, y el costo del uso de la droga para el consumidor bajo la forma de una disminución en la longevidad, la salud, las perspectivas del empleo, etc., tal como ese costo es percibido por el consumidor. La educación acerca de los peligros del consumo de drogas eleva ese costo percibido, pero sobre todo para los consumidores que tienen la cuota suficiente de formación, de preocupaciones por el futuro, de acceso a los medios de comunicación, etc., para entender y responder a los esfuerzos por impartir enseñanza: en suma, los consumidores de clase media. Su demanda de drogas disminuirá, haciendo bajar el precio. La demanda de los consumidores de clase baja disminuirá mucho menos, quizá nada en absoluto, en respuesta a los esfuerzos educativos, y volverá a aumentar a medida que el precio nominal baja debido a que los consumidores de clase media están comprando menos. En último término, es posible que el enfoque educativo sólo desplace la composición de los usuarios más aún hacia los elementos pobres y por lo demás marginales de la sociedad de lo que se da en la situación actual.

Aunque la guerra contra las drogas (haciendo a un lado el costo) se suele considerar un fracaso debido a su efecto menor en el precio de las drogas ilegales y, por lo tanto (suele suponerse) en la demanda, no debemos confundir el costo de un bien con su precio en dólares. El costo para el comprador incluye no sólo el precio, sino también los costos de búsqueda, riesgos para la salud y la posible sanción penal. Como consecuencia de la guerra contra las drogas, se requiere de un poco de esfuerzo para encontrar un vendedor, existe el riesgo de ser arrestado y enjuiciado, y existe el peligro de una sobredosis por accidente a causa de la falta de control de calidad en la fabricación del producto. (También existe un estigma por el uso de drogas ilegales, pero es probable que esto persista si se legalizaran; aunque es legal, el consumo excesivo de alcohol está estigmatizado.) Estos costos se eliminarían si las drogas se legalizaran. Podría parecer que si las drogas valieran más para los consumidores su precio se elevaría. Pero esto es poco probable. El precio estaría limitado al costo por competencia y, por lo tanto, los consumidores obtendrían los beneficios adicionales de las drogas —es decir, los beneficios generados por la eliminación de los costos derivados de la criminalización— como un excedente. Si el bien valiera más para los consumidores pero el precio no fuera mayor, el consumo aumentaría.

VII.12. El castigo penal de los terroristas

Existe un debate acerca de si el terrorismo internacional (principalmente el islamista) debe considerarse como un problema de delincuencia o si más bien

debe juzgarse como una especie de guerra. A diferencia de la postura anterior, se puede argumentar que el derecho penal sólo tiene un efecto disuasivo o incapacitante muy limitado en contra de los terroristas. Incluso si no son terroristas suicidas, es muy difícil disuadirlos debido a la enorme utilidad que obtienen de sus actividades terroristas; y la incapacidad tiene poco efecto porque la oferta de los terroristas parece ser sumamente elástica, ya que las empresas terroristas pueden recurrir a un gran grupo de extremistas musulmanes en todo el mundo. Otro obstáculo para la disuasión es que los delincuentes políticos tienen una expectativa razonable de que serán puestos en libertad mucho antes de que termine la condena de prisión, con frecuencia en un intercambio de rehenes o porque la situación política ha cambiado. Sin embargo, el efecto disuasivo del castigo no es nulo, incluso para los terroristas suicidas; pueden estar dispuestos a morir en un ataque, pero pueden ser muy reacios a pasar muchos años en una prisión si se frustra el ataque.

Los grupos terroristas son más grandes, más sofisticados e incluso tienen mejores finanzas que las bandas delictivas muy grandes; sus recursos pueden exceder a los de los fiscales. Lo que es más importante, el daño potencial causado por el terrorismo (especialmente en una época de proliferación) es a tal punto mayor que el que causan los delitos ordinarios que las protecciones procedimentales otorgadas a los delincuentes comunes pueden imponer costos sociales excesivos. Por ejemplo, la norma que exige la prueba de la culpabilidad más allá de la duda razonable pondera de manera mucho más rigurosa los falsos positivos (la condena de personas inocentes) que los falsos negativos (la exoneración de personas culpables). Esto tiene sentido si el falso positivo es más costoso que el falso negativo, pero puede no ser así si se esperara que un terrorista exonerado por error preparara un ataque terrorista horrible.

El punto más general es que la prevención es un objetivo de la política mucho más importante en el caso del terrorismo mundial que en el caso de la delincuencia común. El país puede vivir con 30 000 asesinatos comunes al año, pero no con 30 000 asesinatos por parte de terroristas. Los castigos penales están diseñados para limitar la tasa delictiva, pero no para reducirla a cero; los costos serían desproporcionados en relación con los beneficios. En el caso del terrorismo esto es mucho menos claro.

La prueba más allá de la duda razonable es sólo un ejemplo de cómo se configura el derecho penal para lidiar con los delitos "comunes". Otro ejemplo es el requisito de que se proporcione al acusado el acceso a las pruebas en su contra, de que su juicio sea público y que tenga derecho a un jurado. En el caso del terrorismo, dar acceso al acusado podría revelar información confidencial sobre los métodos del gobierno para combatir el terrorismo; un juicio público podría proporcionar una plataforma para incitar aún más el terrorismo, y los miembros del jurado podrían ser intimidados por grupos terroristas.

La necesidad de prevenir, junto con las limitaciones del cumplimiento del derecho penal como medio para hacer frente al terrorismo, sugiere que se debe poner mayor atención en la detección de la actividad terrorista inminente que en la investigación posterior a los hechos. A su vez, esto aboga a favor de que se dé al gobierno mayor libertad para llevar a cabo vigilancia electrónica de lo que permite el cumplimiento de la ley, para realizar interrogatorios, para arrestar a personas sospechosas sin la prueba de las actividades terroristas y para castigar la actividad preparatoria y de colaboración (como en las leyes que castigan el "apoyo material" al terrorismo).

VII.13. Las olas delictivas

¿Se pueden explicar en términos económicos las fluctuaciones extraordinarias en la tasa de delincuencia de Estados Unidos? Yo creo que sí.[53]

Los índices delictivos eran muy bajos durante la segunda Guerra Mundial debido a que muchos hombres jóvenes estaban en las fuerzas armadas y prácticamente no existía el desempleo, y se mantuvieron así hasta la década de 1960, cuando se elevaron estratosféricamente, al parecer por razones relacionadas con los disturbios de la década de 1960 y principios de la de 1970, tales como el crecimiento del activismo negro; esto provocó que la policía redujera su presencia en las comunidades negras, teniendo como consecuencia un aumento en los delitos de negros contra negros.

Al principio, la ola delictiva no generó un aumento significativo en encarcelamientos, y esto intensificó la ola. Pero desde el inicio de la década de 1970 y continuando hasta que la recesión económica provocada por la crisis financiera de 2008 golpeó el presupuesto gubernamental, el número de la población reclusa se disparó. En 1950 había sido de 118 por cada 100 000 estadunidenses. De hecho, en 1972 fue menor, 93 por cada 100 000. Pero para el año 2000 había crecido a 469 por cada 100 000. A pesar de que en el decenio de 1990 la tasa de delincuencia en general se redujo en un tercio (y por cerca de dos tercios en muchas ciudades grandes), y el índice de homicidios en más de 40%, la población reclusa continuó creciendo.

Las olas delictivas provocan una reacción política violenta. Tal es el caso de la ola delictiva causada por la inmigración a gran escala de finales del siglo XVIII y principios del XIX, al igual que el de la ola delictiva que inició a principios del decenio de 1960. Durante la presidencia del tribunal de Earl Warren (y en particular durante la década de 1960, cuando los jueces liberales constituían una mayoría sólida), las decisiones liberales de procedimiento penal de la Corte Suprema otorgaron más derechos a presuntos delincuentes y a

[53] El análisis que sigue se basa en William J. Stuntz, *The Collapse of American Criminal Justice* (2011), y en mi reseña de ese libro: "Incarceration Blues", *New Republic*, 36 (17 de noviembre de 2011).

acusados. Esto aumentó los recursos fiscales necesarios para condenar a un acusado (y había más acusados porque había más delincuentes). Hasta que aumentaron los recursos fiscales, creció el porcentaje de delincuentes que o no eran sujetos a proceso o acordaban una sentencia leve con el fiscal, reduciendo el costo del castigo esperado por el delito.

Una Corte Suprema blanda con el delito e indiferente al agudo aumento de la delincuencia al que contribuyeron sus decisiones creó una reacción virulenta en el Congreso y las legislaturas estatales. Esta reacción no sólo se tradujo en un número más alto de policías y de fiscales, sino que también generó una combinación de penas máximas más altas (como la abolición de la libertad condicional), especialmente por delitos federales, para los que las sentencias ya solían ser más altas que para delitos estatales similares, con la proliferación de nuevas prohibiciones penales, muchas de las cuales no requerían de muchas pruebas. El núcleo de este enfoque es viejo; es famosamente ilustrado por la condena de Al Capone por evadir impuestos federales, cuando las fuerzas del orden de Illinois no estaban dispuestas o no tenían la capacidad para atraparlo por sus delitos más graves, que eran violaciones de leyes del estado y no de la ley federal. En la actualidad los fiscales suelen tener varias opciones para aprehender a un delincuente, y la existencia de estas opciones, combinada con sentencias en extremo severas, persuaden a la mayoría de los presuntos delincuentes para declararse culpables, a menudo para cumplir penas rigurosas, en lugar de ir a juicio y recibir una condena aún más rigurosa. Mientras más acusados se declaren culpables, serán más los que se procesen sin ningún aumento en los recursos fiscales. Y, por ende, el régimen de reacción negativa de las penas severas permitió un aumento relativamente barato en el número de condenas, lo que, combinado con sentencias más largas, aumentó tanto la disuasión como la población reclusa (de forma que el gasto penitenciario creció). En la década de 1980 la tasa delictiva se estabilizó en un nivel alto, antes de caer en la década de 1990 (y permanecer a la baja). En 1960, la tasa de homicidios fue de 4.6 por cada 100 000 personas; en 1970, de 7.9; en 1980, de 10.2; en 1990, de 9.4… pero en 2000, de 5.5, y hoy en día es casi la misma, es decir, casi está de vuelta a donde estaba en 1960, antes de que las altas tasas delictivas hicieran del delito un gran problema social y político.

Lecturas recomendadas

Becker, Gary S., "Crime and Punishment: An Economic Approach", 76 *J. Pol. Econ.*, 169 (1968).

————, Kevin M. Murphy y Michael Grossman, "The Market for Illegal Goods: The Case of Drugs", 114 *J. Pol. Econ.*, 38 (2006).

Cooter, Robert, "Prices and Sanctions", 84 *Colum. L. Rev.*, 1523 (1984).

Criminal Law and Economics (Nuno Garoupa [comp.], 2009).

Deis, J. Gregory, "Note, Economics, Causation, and the Entrapment Defense", 2001 *U. Ill. L. Rev.,* 1207 (2001).
Dick, Andrew R., "When Does Organized Crime Pay? A Transaction Cost Analysis", 15 *Int'l Rev. Law & Econ.,* 25 (1995).
Donohue III, John J., y Justin Wolfers, "Estimating the Impact of the Death Penalty on Murder", 11 *Am. Law and Econ. Rev.*, 249 (2009).
Ehrlich, Isaac, "Crime, Punishment and the Market for Offenses", 10 *J. Econ. Perspectives,* 43 (invierno de 1996).
Harel, Alon, "Efficiency and Fairness in Criminal Law: The Case for a Criminal Law Principle of Comparative Fault", 82 *Cal. L. Rev*., 1181 (1994).
Hellman, Daryl A., y Neil O. Alper, *Economics of Crime: Theory and Practice* (5ª ed. 2000).
Levitt, Steven D., y Sudhir Alladi Venkatesh, "An Economic Analysis of a Drug-Selling Gang's Finances", 115 *Q. J. Econ*., 755 (2000).
Murphy, Richard S., y Erin A. O'Hara, "Mistake of Federal Criminal Law: A Study of Coalitions and Costly Information", 5 *Supreme Ct. Econ. Rev.,* 217 (1997).
Parker, Jeffrey S., "The Economics of Mens Rea", 79 *Va. L. Rev*., 741 (1993).
Philipson, Tomas J., y Richard A. Posner, "The Economic Epidemiology of Crime", 39 *J. Law and Econ.*, 405 (1996).
Posner, Richard A., "An Economic Theory of the Criminal Law", 85 *Colum. L. Rev.,* 1193 (1985).
Pyle, D. J., "The Economic Approach to Crime and Punishment", 6 *J. Interdisciplinary Stud.,* 1 (1995).
Schelling, Thomas C., *Choice and Consequences*, cap. 8 (1984).
Schulhofer, Stephen J., "Solving the Drug Enforcement Dilemma: Lessons from Economics", 1994 *U. Chi. Legal Forum,* 207 (1994).
Shavell, Steven, "Criminal Law and the Optimal Use of Non-monetary Sanctions as a Deterrent", 85 *Colum. L. Rev.,* 1232 (1985).
Stigler, George J., "The Optimum Enforcement of Laws", 78 *J. Pol. Econ.,* 526 (1970).

Problemas

1. Evalúe el argumento siguiente, tomado de Stephen F. Williams, "Book Review", 45 *U. Colo. L. Rev*., 437, 450 (1974): "Resulta difícil creer que una multa de 500 dólares tenga el mismo efecto disuasivo sobre nuestro millonario que sobre nuestro empleado postal; presumiblemente, sin embargo, el costo social de una infracción cometida por un hombre rico no es menor que el de la misma infracción cometida por un hombre pobre".

2. Suele afirmarse que los individuos que ocupan posiciones de confianza deben estar bien pagados a fin de disminuir la tentación de traicionar esa confianza. ¿Puede imaginar una base económica para esta concepción?

3. Supongamos que cuesta un dólar de gastos gubernamentales en el cumplimiento de la ley la disminución en 99 centavos de los costos esperados del delito. ¿Estaría tal gasto justificado en términos de costos sociales?

4. ¿Debieran contarse como ganancias sociales las ganancias que obtiene el delincuente por la comisión del delito? (Adviértase que este interrogante es diferente del de si debiera contarse como un costo social el costo de su castigo para el delincuente; todos los que escriben desde una perspectiva económica suponen que sí debiera ser así.) Para centrar la interrogante, supongamos que el derecho contiene una defensa de necesidad suficientemente amplia para volver imposible que se persigan como casos criminales casos tales como el hipotético robo del hombre que se muere de hambre y entra a una cabaña, o quizá aun el de Regina *vs.* Dudley and Stephens. Ahora conteste la pregunta. (Pista: explique en primer término qué diferencia produce la forma como se conteste la pregunta.)

5. ¿Se cometería la violación estatutaria con mayor o menor frecuencia si la joven estuviera sujeta al castigo penal al igual que el hombre? Véase Michael M. *vs.* Sonoma County Superior Court, 450 U.S., 464 (1981).

6. Discuta los costos y beneficios de: dar a los prisioneros adiestramiento en el trabajo; obligarlos a trabajar mientras están en la prisión, y prohibir a los empleadores que se nieguen a contratar a una persona (para un empleo no delicado) porque tiene antecedentes penales.

7. Si la incapacidad mental es autoinducida, como ocurre cuando una persona asesina en un arranque de borrachera, ¿debiera la ley excusar el delito? En caso contrario, ¿debiera castigarlo menos severamente?

8. Supongamos que el demandado no habría cometido el delito del que se le acusa si no fuera porque fue criado en condiciones de pobreza aguda que le impedían adquirir cualesquiera habilidades vendibles en el mercado. ¿Debiera excusarse el delito? ¿Debiera disminuirse, aumentarse o dejarse constante el castigo?

9. Normalmente, el consentimiento no es una defensa contra una acusación penal. Dejando de lado el delito sin víctima, ¿es consistente este delito con el análisis económico?

10. ¿Debiera ser el entrampamiento por parte de un individuo privado una defensa contra una acusación penal en algunas circunstancias?

11. Un homicida en grado de tentativa, aunque no lesione a su presunta víctima ni a nadie más, es castigado de ordinario con mayor severidad que quien consuma un robo. ¿Cuál es la lógica económica de este patrón?

12. Desde el punto de vista de la disuasión, en el caso de que los soldados cometan violaciones de las leyes de la guerra, ¿debiera ser considerado penalmente responsable su comandante: *a)* si se entera de las violaciones pero no hace nada para impedirlas, *b)* independientemente de que se entere o no se entere, o *c)* sólo si ordena o alienta las violaciones? (Véase más adelante XV.15).

13. Analice las consecuencias económicas de los programas en los que el gobierno ofrece comprar de vuelta las armas cortas de sus propietarios, en un esfuerzo por reducir los delitos de armas. Véase Levitt, en VII.2, nota 4, p. 307; Wallace P. Mullin, "Will Gun Buyback Programs Increase the Quantity of Guns?", 21 *Int'l Rev. Law and Econ.*, 87 (2001).

14. ¿Son las películas violentas sustitutos o complementos de los delitos violentos? Véase Gordon Dahl y Stefano DellaVigna, "Does Movie Violence Increase Violent Crime?", 124 *Q. J. Econ.,* 677 (2009).

VIII. EL DERECHO COMÚN, LA HISTORIA DEL DERECHO Y LA JURISPRUDENCIA

Aunque no está completo aún nuestro estudio económico del derecho común —en particular, el capítulo xix examinará el derecho de los fideicomisos y el derecho patrimonial, y el capítulo xxii el derecho común del procedimiento civil y penal, incluido el conflicto de leyes—, hemos avanzado lo suficiente para hacer una consideración del derecho común en conjunto, uniendo algunas de las ideas desarrolladas en los capítulos anteriores y construyendo a partir de ellas lo viable y lo útil. Este capítulo considera también algunos temas que se cubren normalmente en los cursos de la escuela de derecho sobre la jurisprudencia, la historia del derecho y la antropología jurídica.

VIII.1. La lógica económica implícita del derecho común

Para la mayoría de los abogados, el derecho común es una colección de campos desiguales, cada uno de ellos con su propia historia, vocabulario, y una profusión desconcertante de reglas y doctrinas; en efecto, cada campo podría parecer una colección de doctrinas apenas tenuemente relacionadas. Pero hemos visto que el derecho de la propiedad, el derecho contractual, el derecho mercantil, el derecho aplicable a los cuasidelitos y de la restitución y el enriquecimiento injusto, el derecho penal y familiar y el derecho marítimo pueden enunciarse en términos económicos que no sólo explican las doctrinas principales, tanto sustantivas como remediales, de estos campos del derecho elaborado (en gran medida) por los jueces, sino que también prueban su unidad económica.[1] Otra manera de plantear esto (aunque con cierta exageración) es que la economía es la estructura profunda del derecho común y las doctrinas de este derecho son la estructura superficial. En los contextos donde es bajo el costo de las transacciones voluntarias, las doctrinas del derecho común crean incentivos para que la gente canalice sus transacciones a través del mercado (ya sea implícito —el mercado matrimonial, por ejemplo— o explícito). Lo hacen así creando derechos de propiedad (ampliamente definidos)

[1] Recuérdese que, como se señaló al principio del capítulo iii, "derecho común" se emplea a lo largo de este libro en un sentido no técnico, como un sinónimo del derecho elaborado primordialmente por los jueces antes que por los legisladores o los creadores de constituciones. En términos técnicos, campos tales como el derecho familiar, el marítimo, y aun el de la equidad, aunque son elaborados por jueces, no constituyen campos del derecho común porque fueron elaborados por jueces de tribunales que no se consideraban tribunales del derecho común.

y aplicándolos mediante remedios diseñados para impedir las transferencias coercitivas —remedios tales como los requerimientos, la restitución, los daños punitivos y el castigo penal—. En los contextos donde el costo de la asignación de los recursos mediante transacciones voluntarias es prohibitivamente elevado, convirtiendo al mercado en un método inviable para la asignación de los recursos, el derecho común fija precios para el comportamiento en una forma tal que se imite al mercado. Por ejemplo, el sistema de los cuasidelitos asigna la responsabilidad de los accidentes entre el ferrocarril y el agricultor, el conductor y el peatón, el médico y el paciente inconsciente con el que el médico se topa y al que trata sin un contrato previo, en forma tal que se produzca la asignación de recursos aproximada para la seguridad que generaría el mercado si se le pudiera hacer funcionar. El derecho contractual hace lo mismo en lo tocante a las contingencias que podrían imposibilitar la realización de un contrato: impone la responsabilidad a la parte que sea más capaz de impedir que ocurra la contingencia o de minimizar la desutilidad de su ocurrencia comprando un seguro o autoasegurándose. El derecho de la propiedad hace lo mismo limitando los derechos de propiedad en situaciones donde la insistencia en un derecho absoluto impediría un intercambio maximizador del valor. Invirtiendo el orden de discusión anterior, digamos que el derecho común establece derechos de propiedad, regula su intercambio y los protege contra una interferencia irrazonable: todo ello con el fin de facilitar la operación del mercado libre, así como, cuando el mercado libre no puede funcionar, con el de simular sus resultados.

De acuerdo con la metáfora de la "estructura profunda", el analista económico del derecho puede moverse con facilidad entre los campos del derecho común. Por ejemplo, casi cualquier problema de cuasidelitos puede replantearse y resolverse como un problema contractual, tan sólo preguntándose cómo habrían convenido por adelantado las personas involucradas en un accidente lo tocante a las medidas de seguridad si los costos de transacción no hubieran sido prohibitivos. Consideremos el caso de Eckert *vs.* Long Island Railroad.[2] El tren de la parte demandada se desplazaba a exceso de velocidad y sin señales adecuadas en un área densamente poblada. Un niño pequeño estaba sentado en las vías sin prestar atención al tren que se aproximaba. Eckert corrió para rescatar al niño y pudo sacarlo de las vías pero él mismo resultó muerto. El tribunal sostuvo que Eckert no había sido contribuyentemente negligente, de modo que sus herederos podrían recuperar los daños causados por la negligencia del ferrocarril. Porque "no era indebido de su parte que hiciera todo lo que estuviera en sus manos para rescatar al niño, algo que fuera compatible con una preocupación razonable por su propia seguridad. Era su obligación ejercitar su criterio acerca de que pudiera salvar probablemente al niño sin lesión grave para él mismo".[3] Si, como lo implica este pasaje, la

[2] 43 N.Y., 502 (1870).
[3] *Ibid.*, 505.

probabilidad de que el niño hubiese muerto si no se hubiera intentado el rescate era mayor que la probabilidad de que Eckert hubiera muerto al salvar al niño, y si la vida del niño era por lo menos tan valiosa como la de Eckert, el beneficio esperado del rescate para el ferrocarril al disminuir un costo esperado de la responsabilidad para con los padres del niño era mayor que el costo esperado del rescate. En ese caso, si no hubiese sido por los costos de transacción prohibitivos, el ferrocarril habría contratado a Eckert para que intentara el rescate, de modo que se le habría exigido que compensara a los herederos de su patrimonio *ex post*.[4]

Igualmente, casi cualquier problema contractual puede resolverse como un problema de cuasidelito preguntándonos cuál sanción es necesaria para impedir que la parte involucrada en un contrato observe una conducta socialmente dispendiosa, tal como aprovecharse de la vulnerabilidad de una parte que cumple primero lo que le toca en la negociación. Y los problemas de cuasidelitos y de contratos pueden presentarse como problemas en la definición de los derechos de propiedad; por ejemplo, podría pensarse que el derecho de la negligencia define el derecho de "propiedad" que tenemos a la seguridad de nuestras personas contra lesiones accidentales. La propia definición de los derechos de propiedad puede verse como un proceso de determinación de las medidas en que habrían convenido las partes si los costos de transacción no fuesen prohibitivos, a fin de crear incentivos para evitar el desperdicio de recursos valiosos.

Las distinciones importantes en el análisis económico del derecho común atraviesan las diversas divisiones convencionales de los temas. Una se establece entre los casos en los que sólo se requiere la compensación si se ha dejado de tomar alguna medida para evitar la pérdida que minimice el costo y los casos en que se requiere la compensación de todos modos. Quien incumple un contrato debe pagar de ordinario daños, aunque su incumplimiento conduzca a un uso más valioso de los recursos; lo mismo se aplica a un allanador. Pero alguien que accidentalmente se inflija un perjuicio personal en circunstancias en las que ninguna precaución justificada por el costo pudiera haber impedido el accidente en general no es responsable; y en la mayoría de los casos de responsabilidad penal se requiere la intención de causar daño. Hay razones económicas para estas diferencias. Los elevados costos de las sanciones penales aconsejan que se las limite a los casos donde es pequeño el riesgo del error legal. La regla de la compensación en casos contractuales es un corolario de hacer de la responsabilidad estricta el estándar por el incumplimiento de nuestras promesas contractuales, que según hemos visto es económicamente justificable, en parte debido a la función de seguro de muchas promesas contractuales. La regla de la responsabilidad estricta en los

[4] ¿Qué pasa si la probabilidad de que Eckert muera al intentar el rescate es muy grande? En términos de la Fórmula de Hand, eso podría elevar a B a un punto en el que no habría beneficio neto alguno para el ferrocarril si Eckert rescatara al niño.

casos de allanamiento, pero no en los casos del accidente ordinario, se basa en la distinción existente entre los casos donde los costos de transacción son muy elevados (la gran mayoría de los casos de accidentes) y los casos donde, porque hay una relación efectiva o potencial de comprador-vendedor entre las partes de la interacción (el allanador podría negociar un permiso para utilizar la tierra que es propiedad de otra persona) y se satisfacen otras condiciones (información adecuada, capacidad mental de las partes, ausencia de monopolio —incluido el monopolio bilateral—, un problema recurrente en pocos casos, según hemos visto), los costos de transacción son relativamente bajos y las transacciones voluntarias debieran alentarse en consecuencia (véase, por ejemplo, III.4).

No debiera creerse que la última distinción implique que la asignación legal de los derechos o las obligaciones carezca de importancia en los casos donde los costos de transacción son bajos. La corrección de las externalidades, tanto positivas (beneficios externos) como negativas (costos externos), es sólo una función económica del derecho: éste también tiene una función en la disminución de los costos de transacción y en hacer posibles o facilitar, por oposición a la simulación, los procesos del mercado. Las reglas que llenan los vacíos en la interpretación de contratos son un ejemplo de esta función; reducen los costos de negociación y de elaboración. Otro ejemplo es la susceptibilidad del derecho, que se muestra ante el monopolio bilateral, por ejemplo, en la doctrina marítima sobre el rescate. El monopolio bilateral aumenta los costos de transacción, a veces hasta el punto en que se vuelven imposibles las transacciones que maximizan el valor. Pero, aun cuando la transacción ocurre, lo hace a un costo mayor que si no hubiera ningún monopolio bilateral. Es probable que, ante la presencia de un monopolio bilateral, el tribunal trate de adivinar el acuerdo que habrían negociado las partes si las negociaciones hubieran sido factibles a un precio razonable, antes que tratar de obligarlas a negociar.

Si el derecho común es una unidad económica, entonces debiéramos esperar que los casos económicamente similares se decidan en la misma forma aunque surjan en campos diferentes. Un ejemplo será ilustrativo.[5] A le compra una impresora a B. Éste contrata a C para que transporte la máquina a las instalaciones de A. C realiza la tarea descuidadamente y, en consecuencia, le cuesta a A un tiempo inesperadamente largo la instalación de la impresora y durante este periodo sufre una pérdida de ganancias, por lo que demanda a C por un cuasidelito, acusándolo de negligencia. Si se tratara de un juicio contractual, la regla de Hadley *vs.* Baxendale impediría que A obtuviera sus ganancias perdidas, pero es un juicio de cuasidelito; A no tiene ningún contrato con C. Pese a ello, A no podrá obtener las ganancias perdidas. La razón es que C no podría haber determinado fácilmente el impacto de la de-

[5] Véase Rardin *vs.* T & D Machine Handling, Inc., 890 F.2d, 24 (7th Cir., 1989).

mora sobre el negocio de A —ellos no se conocían—, mientras que A podría haberse protegido negociando una cláusula de daños líquidos con B, quien podría haber seleccionado entonces a un contratista diferente o informarle a C de la necesidad de un cumplimiento pleno de su compromiso.

Empero, el lector podría preguntarse cómo es posible que el derecho común —un cuerpo antiguo de doctrina jurídica, que sólo ha cambiado marginalmente en el último siglo— tenga tanto sentido económico como parece hacerlo. Una respuesta completa (hasta donde es posible) deberá esperar hasta los capítulos XX a XXII, donde se discutirán los procedimientos y el ambiente institucional del derecho común. Pero aquí tenemos algunas sugerencias.

1. Muchas doctrinas del derecho común son económicamente razonables pero no económicamente sutiles. Son de sentido común. Su articulación en términos económicos escapa a la capacidad de la mayoría de los jueces y los abogados, pero no a su intuición.
2. Lo que Adam Smith llamaba la riqueza de una nación, lo que aquí llamamos eficiencia y lo que un lego podría llamar el tamaño del pastel, ha sido siempre un valor social importante… y nunca tanto como en el siglo XIX, la era de los principios del *laissez-faire* (mercado libre), cuando el derecho común adquirió gran parte de su forma moderna. No es sorprendente que este valor haya influido en las decisiones judiciales, en especial desde que estos principios han hecho una reaparición extraordinaria en el pensamiento estadunidense y del mundo entero en los últimos 25 años.
3. Los valores sociales que compiten con la eficacia son más controversiales y difíciles de implementar con las limitadas herramientas de que disponen los jueces para realizar su trabajo. Esos valores se refieren sobre todo a las ideas referentes a la distribución justa del ingreso y la riqueza, conceptos sobre los que no se ha generado un consenso. De cualquier forma, las políticas redistributivas eficaces requieren facultades de tributación y gasto de las que carecen los jueces (véase por ejemplo, después, XVII.7). Dado que no pueden hacer mucho como jueces del derecho común para alterar la medida de las rebanadas del pastel que reciben los diversos grupos de la sociedad, también podrían concentrarse en el incremento de su tamaño.
4. El hecho de que, con pocas excepciones, los abogados y los jueces no sean deliberadamente económicos en su enfoque del derecho es una objeción trivial contra el análisis económico positivo del derecho común. El lenguaje de la economía es un lenguaje diseñado para los académicos y los estudiantes, no para las personas cuyo comportamiento estudia el economista. Los poetas no emplean el vocabulario de los críticos literarios; los jueces no emplean el vocabulario de los economistas.

A pesar de todo lo anterior, no todas las doctrinas del derecho común tienen una justificación económica sólida. Entre las excepciones se encuentran el método para el cálculo de los daños en los casos de muerte (y la negativa del derecho común a otorgar el pago de daño alguno por la pérdida de la vida; tales pagos son una innovación de las leyes reguladoras de mediados del siglo XIX) y la sustitución de la negligencia contribuyente por la negligencia comparativa, y de la no contribución por la contribución entre los cuasidelincuentes conjuntos.

Algunas de las discrepancias existentes entre el derecho y la economía podrían deberse simplemente a las demoras explicables en términos económicos, el fenómeno que los economistas llaman "dependencia de la ruta" (véase después XXI.4). Debido a que otorga un gran peso a la continuidad (y hay buenas razones económicas y políticas para ello), el derecho tiende a ir detrás del cambio de las condiciones sociales y económicas. La literatura reciente considera que las diferencias en las leyes modernas de los distintos países están determinadas principalmente por factores como, por ejemplo, si el ordenamiento jurídico de un país desciende del derecho común inglés o del derecho francés, incluso cuando los orígenes se remontan a cientos de años atrás (véase después XXXI.1).

VIII.2. Los cuasidelitos económicos

Mencioné la superposición entre los campos del derecho común. La superposición entre los cuasidelitos y los contratos la ilustra bien lo que se conoce como "cuasidelitos económicos". La doctrina de la pérdida económica (véase antes VI.7) aísla la responsabilidad por causar una pérdida económica sin ocasionar daño alguno a la propiedad o lesiones personales. Desde luego, el daño a la propiedad y las lesiones personales también son pérdidas "económicas", pero el concepto legal de los "cuasidelitos económicos" (un término que incluye acciones dañinas interpretadas como no procesables por parte de la doctrina de la pérdida económica) los excluye.

Es un enigma la razón por la que la doctrina de la pérdida económica ya no se aplique en los casos de tergiversación en la mayoría de las jurisdicciones: la causa de que, en definitiva, la opinión de Cardozo en el caso Ultramares[6] —el origen del principio de que un auditor que tergiversa la situación financiera de la empresa que audita no es responsable ante los prestamistas de la empresa por las consecuencias de la tergiversación, a pesar de que ellos confiaron sensatamente en su auditoría para decidirse a prestar a la empresa— ahora sea minoritaria. Es cierto que, en dicho caso, las víctimas están mejor delimitadas que en muchos casos de pérdida económica: son los pres-

[6] Ultramares Corp. *vs.* Touche, 174 N.E. 441 (N.Y., 1931).

tamistas de la empresa (o en algunos casos sus otros acreedores, o sus accionistas) lo que representa un asunto de interés público. Y es cierto que el auditor tendrá una mejor idea de la magnitud potencial de una tergiversación, ya que sabe mucho acerca de la empresa que está auditando. Sin embargo, la diferencia entre el costo privado y el costo social, que como el lector recordará es la justificación principal de la doctrina de la pérdida económica, también es mucho mayor en el caso del auditor. Su tergiversación consistirá normalmente en ocultar las malas noticias sobre la empresa objeto de la auditoría. Con el tiempo, esas malas noticias se filtrarán o de otro modo saldrán a la luz, por lo que su ocultamiento habrá transferido riqueza de un grupo de prestamistas o accionistas a otro, antes que haber provocado una disminución neta de la riqueza total de los prestamistas o de los accionistas. La pérdida económica neta será la inyección de una incertidumbre adicional en los mercados crediticios o de valores, haciendo que esos mercados sean un poco menos eficientes. Ésta es una pérdida distinta y menor a las pérdidas privadas de los prestamistas que decidieron prestar o no prestar a la empresa, o de los accionistas a quienes se persuadió para comprar o no vender sus acciones.

Ultramares pertenece a un tipo de casos en los que el incumplimiento de contrato (a diferencia del cuasidelito, como en el caso Rickards, véase antes VI.7) perjudica a un tercero. El más simple de estos casos es aquel en el que el incumplimiento tiene como consecuencia una transferencia de riqueza incorrecta, como en el caso White *vs.* Jones (analizado junto con Rickards), en el que un abogado, por negligencia, no incluyó una herencia para las hijas del testador en el testamento que redactó para éste, y se le consideró como responsable del cuasidelito por este error. ¿Podría, en cambio, haber sido demandado por incumplimiento de contrato, según la teoría de que las hijas eran terceros beneficiarios del contrato por el cual se contrató al abogado para redactar el testamento? Una persona es un tercero beneficiario si las partes involucradas en el contrato tenían planeado que tuviera los derechos de una parte, específicamente para demandar en el caso de un incumplimiento y recuperar los daños causados a ella por el incumplimiento. Considerar a los legatarios previstos como terceros beneficiarios de un contrato entre el testador y su abogado especialista en derecho hereditario o patrimonial expondría al abogado a una mayor responsabilidad y éste cobraría más por preparar el testamento. El cliente podría decidir si vale la pena pagar honorarios más elevados a fin de reducir el riesgo (al hacer del legatario previsto un tercero beneficiario y, por lo tanto, un responsable potencial del cumplimiento del testamento) de que herede el legatario equivocado.

Restructurar el cuasidelito como un contrato no dará resultado cuando la realización negligente de un contrato entre un empleador y un tercero tenga consecuencias adversas para el empleado. Supongamos que el empleador solicita a un empleado que se realice una prueba de orina para detectar la presencia de drogas ilegales en su cuerpo y el laboratorio al que se envió la

muestra de orina reporta por error que el empleado dio positivo. Es poco probable que las partes involucradas en el contrato —el empleador y el laboratorio— tengan la intención de que el empleado pueda hacer cumplir su contrato, el cual demandaba de manera expresa o implícita que el laboratorio tomara el debido cuidado en la evaluación de los resultados de la prueba; de modo que el empleado no es un tercero beneficiario. ¿Debiera permitírsele presentar una demanda de cuasidelito en contra del laboratorio? Los tribunales están divididos respecto a este tema. El argumento en contra de la responsabilidad es que el empleador tiene un incentivo válido para hacer cumplir el contrato, y si se descubre el resultado erróneo lo dará por terminado o demandará mayor cuidado o un precio más bajo y volverá a contratar al empleado despedido por error. Sin embargo, esto resulta inverosímil. Por lo general, los empleadores sólo solicitan a los empleados de nivel inferior que se sometan a pruebas de dopaje, y un falso positivo esporádico perjudicará poco al empleador; le será fácil encontrar trabajadores sustitutos. La víctima principal y, por lo tanto, el responsable lógico que haga cumplir la obligación del laboratorio de tomar el cuidado debido es el empleado.

Otro ejemplo de la restructuración del contrato como un cuasidelito es el cuasidelito que constituye el daño ocasionado por la interferencia intencional con un contrato (o, lo cual no requiere de una discusión aparte, la interferencia intencional con las relaciones comerciales ventajosas que aún no se han materializado en un contrato).[7] Si A y B tienen un contrato vigente, y C persuade a B para romperlo y que lo recontrate a él (C) en su lugar, C ha cometido el cuasidelito. A simple vista, el resultado parece opuesto a los preceptos de la eficiencia y, en especial, a la doctrina del incumplimiento eficiente. Supongamos que el contrato había especificado que A pagaría 100 000 a B por el cumplimiento de éste y que si B rompía el contrato A sufriría daños por 10 000. Entonces B no estaría de acuerdo en romper su contrato con A y recontratar con C a menos que C se comprometiera a pagarle a B al menos 110 000 (100 000 por el cumplimiento en sí y 10 000 para compensar a A por el incumplimiento de contrato de B). C no estaría de acuerdo a no ser que el cumplimiento de B valiera por lo menos 110 000 para él, de modo que ése sería el precio mínimo en el nuevo contrato, el contrato entre B y C. Al final de la secuencia del incumplimiento y la recontratación, A no estaría peor ni mejor que antes; B no estaría peor —porque al menos tendría 110 000, y podría estar mejor si C le pagara más— y C estaría mejor por la diferencia entre lo que había acordado pagarle a B y lo que valía para él el cumplimiento de B. Una parte y quizá dos estarían mejor y ninguna estaría peor, de modo que la secuencia es eficiente, ¿y entonces por qué C debiera ser culpable de un cuasidelito? Desde luego, si es disuadido por la amenaza de la responsabilidad civil de hacer trato directamente con B —si, en cambio, se le fuerza a arre-

[7] Compárese William M. Landes y Richard A. Posner, *The Economic Structure of Tort Law*, 222-225 (1987), con Lillian R. BeVier, "Reconsidering Inducement", 76 *Va. L. Rev.*, 877 (1990).

glar un acuerdo de tres partes en donde A pudiera interrumpir voluntariamente su contrato con B—, C podría lograr el mismo resultado, pero el costo sería más alto debido a la complejidad de la negociación.

Pero supongamos que B es insolvente. Ya no hay motivos para confiar en que el incumplimiento y la recontratación producirían una ganancia neta en el bienestar económico; la pérdida no compensada de A podría superar el beneficio de B y C. O supongamos que no hay recontratación sino tan sólo un incumplimiento provocado por la información errónea transmitida por C a B; entonces C debiera ser responsable de las consecuencias de haberla obtenido. En este caso, ninguna de las apreciaciones que apoya la doctrina de la pérdida económica debate en contra de la responsabilidad por un cuasidelito.

En ocasiones, el incumplimiento de mala fe de un contrato se reclasifica como un cuasidelito con el fin de otorgar el derecho a la víctima del incumplimiento a demandar por daños punitivos, los cuales no se conceden en una demanda por incumplimiento de contrato. El ejemplo principal es la negativa de mala fe de una compañía de seguros a cumplir con el contrato de seguro (véase antes IV.12). Otra táctica evasiva en torno al derecho de contractual que se discutió en esa sección es que se permita que un paciente demande a su médico por el cuasidelito de negligencia profesional antes que sólo demandar por incumplimiento de contrato entre el demandante y él.

VIII.3. El derecho común, el crecimiento económico y la historia del derecho

El análisis económico puede ayudar a aclarar el papel controversial desempeñado por el derecho común en el crecimiento económico (el aumento a lo largo del tiempo en la producción de una sociedad) de los Estados Unidos. La opinión general es que el derecho común ayudó a promover el desarrollo económico en el siglo XIX al adoptar una postura permisiva, incluso facilitadora, hacia la actividad empresarial. Una variante es que subsidió el crecimiento al no hacer que la industria soportara todos los costos que habría requerido un compromiso genuino con la eficiencia. La permisividad del derecho común en el siglo XIX se contrasta con las numerosas restricciones impuestas a la actividad económica por el derecho en el periodo precedente y el siguiente.[8]

Dado que el crecimiento se impulsa cuando se emplean los recursos eficientemente, hay un sentido, aunque algo controversial, en el que podría decirse que el derecho común, en la medida en que se ha forjado por una preocupación por la eficiencia, ha promovido el crecimiento. Sin embargo, la sociedad puede acelerar el ritmo del crecimiento obligando a la gente a consumir menos

[8] Ésta es la tesis de Morton J. Horwitz, *The Transformation of American Law, 1780-1860* (1977). Véanse las críticas formuladas desde un punto de vista económico en Stephen F. Williams, "Book Review", 25 *UCLA L. Rev.,* 1187 (1978); Herbert Hovenkamp, "The Economics of Legal History", 67 *Minn. L. Rev.,* 645, 670-689 (1983).

y ahorrar más, y aumentando los rendimientos de la inversión de capital. Si el derecho común desempeñó algún papel en la aceleración del crecimiento económico, lo debe haber hecho volviendo más rentable la inversión de capital.

En este tenor, se ha sostenido que el derecho contractual del siglo XIX favoreció consistentemente a la parte desempeñante sobre la pagante a fin de estimular el espíritu empresarial.[9] El argumento no es convincente. Toda empresa es simultáneamente, y más o menos igualmente, desempeñante y pagante. La empresa es desempeñante respecto de los contratos de venta de su producción; es el pagante respecto de los contratos de compra de sus insumos. No obtiene ninguna ganancia clara del hecho de tener al derecho inclinado a favor de los desempeñantes.

Se ha sugerido que el derecho común de los accidentes industriales favorecía a la industria.[10] Pero, al igual que entre las partes que ya mantienen una relación contractual, podría alcanzarse el nivel eficiente de la seguridad, aunque el derecho no impusiera *ninguna* responsabilidad por el daño accidental. Esto sería cierto aunque los salarios se encontraran al nivel de subsistencia, el nivel al que cualquier reducción adicional de los salarios haría que el trabajador se muriera de hambre o se debilitara tanto que no podría trabajar con eficacia. A un nivel de subsistencia, los trabajadores se negarían a intercambiar una reducción salarial por un aumento de la seguridad del trabajo, pero ésta sería la decisión óptima para ellos: es preferible asumir algunos riesgos que morirse de hambre. No obstante, aun en los Estados Unidos del siglo XIX los salarios industriales estaban muy por encima de los niveles de subsistencia (ésa es una razón para que hubiera una inmigración tan grande).[11] De modo que debió de haber competencia por los trabajadores, y una forma de competir por los trabajadores consiste en ofrecer condiciones de trabajo más seguras. Los costos de la publicidad de un trabajo más seguro pueden ser elevados, sobre todo cuando muchos trabajadores no hablan el idioma; pero la ignorancia de los peligros debe de haber sido igualmente generalizada en la Inglaterra del siglo XVIII, y sin embargo sabemos que los trabajadores de empleos peligrosos o desagradables recibían entonces considerables primas salariales.[12] Ahora, muchos peligros potenciales del lugar de trabajo son tan suti-

[9] Grant Gilmore, "Products Liability: A Commentary", 38 *U. Chi. L. Rev.,* 103 (1970).

[10] Por ejemplo por el profesor Horwitz, nota 8, arriba, en 97-101. Véanse opiniones contrastantes sobre este punto en Richard A. Posner, "A Theory of Negligence", 1 *J. Legal Stud.,* 29, 67-72 (1972), y Gary T. Schwartz, "Tort Law and the Economy in the Nineteenth-Century America: A Reinterpretation", 90 *Yale L. J.,* 1717 (1981). Aunque critica el derecho de los accidentes industriales, el profesor Schwartz rechaza la teoría del siglo XIX sobre el subsidio en el derecho aplicable a los cuasidelitos.

[11] Jeffrey G. Williamson, *Late Nineteenth-Century American Development: A General Equilibrium History* 240-243, 249 (1974); Maurice Wilkinson, "European Migration to the United States: An Econometric Analysis of Aggregate Labor Supply and Demand", 52 *Rev. Econ. & Stat.,* 272 (1970).

[12] Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations,* 112-113, 116-

les que su evaluación resultaría difícil o aun imposible para los trabajadores. Esto era mucho menos cierto en el siglo XIX, cuando los peligros sutiles escapaban probablemente a la atención de todos.

El derecho común de los accidentes industriales fue sustituido en la mayoría de las industrias, a principios del siglo XX, por el sistema conocido actualmente como compensación de los trabajadores (llamado originalmente compensación de los "obreros"), el que vuelve al empleador estrictamente responsable de las lesiones resultantes de los accidentes industriales, pero provee el pago de daños fijos, limitados, en lugar de la evaluación individualizada del derecho aplicable a los cuasidelitos. En virtud de que los pagos de compensación de los trabajadores son menores de lo que serían los pagos en los casos del derecho común, disminuye el incentivo del empleador para tener cuidado, mientras que el del trabajador aumenta. (Este último efecto es similar al de los deducibles de las pólizas de seguros.[13]) En consecuencia, el efecto neto sobre la tasa de accidentes difiere de una industria a otra.[14] Otra complicación es el hecho de que el costo de litigar la compensación para los trabajadores es menor que el costo de litigar el derecho común de los cuasidelitos;[15] los intereses son menores porque la compensación, al fijarse mediante un tabulador en lugar de calcularse de nuevo en cada caso, es menor. Y no existe cuestión de responsabilidad alguna que deba juzgarse. Los costos menores de litigación beneficiarían a los trabajadores principalmente; sin embargo, el menor costo de la compensación de los trabajadores explica por qué los patrones no se resistieran vigorosamente a su sustitución en lugar de la responsabilidad del derecho común.[16] Recordemos que éste es el tipo de intercambio propuesto por los defensores de la no culpabilidad del seguro de accidentes de automóvil (véase antes VI.14).

Es más probable que la responsabilidad estricta de los accidentes industriales tenga sentido si los costos de la información sobre seguridad en el trabajo son altos. Puesto que, si los costos de información impiden la autoprotección óptima por parte del empleado, los únicos cambios en el cuidado o en la actividad que puedan optimizar los accidentes serán los cambios que

117 (Edwin Cannan [comp.], 1976) [edición en español: *Una investigación sobre la naturaleza y causas de la riqueza de las naciones,* trad. de Gabriel Franco, Fondo de Cultura Económica, México, 1982]. Por lo que toca al comportamiento de los trabajadores modernos en este respecto, véase W. Kip Viscusi, "Wealth Effects and Earning Premiums for Job Hazards", 60 *Rev. Econ. & Stat.,* 408 (1978), y los estudios allí citados; W. Kip Viscusi, *Risk by Choice: Regulating Health and Safety in the Workplace*, cap. 3 (1983); y las referencias de VI.12 antes, nota 61.

[13] Véase en general *Safety and the Work Force: Incentives and Disincentives in Workers' Compensation* (John D. Worrall [comp.], 1983).

[14] Véase Price V. Fishback y Shawn Everett Kantor, *A Prelude to the Welfare State: The Origins of Workers' Compensation,* 12, 77-83 (2000); Alan B. Krueger, "Incentive Effects of Workers' Compensation Insurance", 41 *J. Pub. Econ.,* 73 (1990).

[15] Como lo destacan Fishback y Kantor, *op. cit.,* en 11.

[16] Véase *idem.*

realice el empleador. Éste es el argumento más fuerte a favor de la superioridad de la compensación del trabajador sobre la responsabilidad del derecho común de los accidentes industriales. El movimiento paralelo en la responsabilidad por los productos, desde virtualmente ninguna responsabilidad en el siglo XIX hasta la responsabilidad semiestricta de la actualidad (véase antes VI.6), podría estar relacionado también con los crecientes costos de información para víctimas potenciales.

El argumento de que el derecho común del siglo XIX favorecía al crecimiento es ligeramente más plausible en lo tocante a accidentes de extraños. Consideremos dos reglas de derecho, una de las cuales establece que un ferrocarril es responsable ante los transeúntes lesionados en los cruces ferroviarios sólo si el ferrocarril fue negligente, mientras que la otra establece que el ferrocarril es estrictamente responsable, a menos, quizá, que los transeúntes fuesen contribuyentemente negligentes. La tasa de accidentes será similar con ambas reglas, pero los costos del ferrocarril serán mayores con la segunda regla, lo que conducirá a un aumento de sus precios y una disminución de su producción y sus ganancias. La primera regla alienta y la segunda desalienta a los ferrocarriles, aunque quizá trivialmente.

Sin embargo, no se concluye de aquí que la primera regla constituya un subsidio para los ferrocarriles en ningún sentido peyorativo del término. Como vimos en el capítulo VI, la elección del planificador social neutral entre la negligencia y la responsabilidad estricta implica un cálculo complejo, y no está claro que el cálculo apunte hacia una responsabilidad estricta para los accidentes ferroviarios. Y aunque lo hiciera, de modo que pudiera afirmarse que la elección de la negligencia había sido un subsidio para los ferrocarriles, es posible que haya habido una justificación económica. En el siglo XIX, la construcción de una línea de ferrocarril aumentaba el valor de la tierra a lo largo de la línea, otorgando a los terratenientes mejor acceso a los mercados. Pero, a menos que el ferrocarril fuera propietario de la tierra afectada, o pudiera negociar antes de la construcción con cada terrateniente, no podría captar todo el aumento del valor y, por lo tanto, no invertiría en nuevas construcciones hasta el punto en que el último dólar gastado aumentara los valores de las tierras[17] en un dólar exactamente. Si los tribunales escogieron la menos costosa (para los ferrocarriles) de las dos reglas de la responsabilidad eficiente porque percibían que los ingresos de los ferrocarriles eran menores de lo que debieran ser para estimular el nivel económicamente correcto de la inversión

[17] Valores *netos*: toda reducción del valor de la tierra debida a la construcción de la línea (por ejemplo, la línea podría reducir la ventaja de localización de la tierra más cercana a los mercados) debería restarse de los aumentos. Véanse las estimaciones de los beneficios sociales de los ferrocarriles en los Estados Unidos del siglo XIX en Robert William Fogel, "Railroads and American Economic Growth", en *The Reinterpretation of American Economic History,* 187 (Robert William Fogel y Stanley L. Engerman [comps.], 1971); Patrick O'Brien, *The New Economic History of the Railways* (1977).

en la industria ferroviaria, es posible que hayan aproximado a la economía un poco más al empleo óptimo de sus recursos. Vimos en el capítulo VI que la externalización de los costos de los accidentes es un método para estimular la provisión de beneficios externos.

En esta conexión debemos tomar en cuenta la inmunidad tradicional de las empresas caritativas frente a la responsabilidad de los cuasidelitos. Como veremos en el capítulo XVII, un problema de "gorroneo" (beneficio externo) impide que el mercado provea el nivel eficiente de los servicios caritativos. Un procedimiento para reducir la gravedad de este problema consiste en permitir que las instituciones de caridad externalicen algunos de sus costos. Sin embargo, virtualmente todos los estados de la Unión Americana han abolido la inmunidad de las instituciones de caridad en lo tocante a los cuasidelitos, quizá porque la caridad es un bien superior: un bien cuya demanda aumenta más que proporcionalmente con el aumento de los ingresos. Por lo tanto, el aumento de los ingresos reales se traduciría en un aumento más que proporcional en donaciones de caridad, por lo que es menos necesario permitir que las instituciones de caridad reduzcan sus costos volviéndose inmunes a la responsabilidad de los cuasidelitos. Además, la deducción de las caridades del impuesto sobre la renta provee un método más eficiente para estimular la provisión de servicios caritativos (véase después XVIII.8). La inmunidad favorece de manera arbitraria a las instituciones de caridad que se dedican a actividades que pueden dar lugar a demandas de responsabilidad o de cuasidelitos; la deducción fiscal no lo hace.

Un crítico, el profesor Horwitz, ofrece el siguiente "ejemplo importante [en el derecho común] del sesgo de clase":[18] en un contrato de trabajo por un periodo convenido (usualmente un año) con salarios pagaderos al final del periodo, un empleado que incumpliera el contrato no tendría derecho a recuperar el valor de su trabajo realizado hasta el momento del incumplimiento (menos cualesquiera daños causados al empleador); pero si un constructor incumpliera un contrato de construcción sin falta de su parte, tendría derecho a tal recuperación. La crítica pasa por alto una distinción obvia: postergar el pago en el caso del contrato de trabajo era un instrumento para asegurar que el empleado completara el periodo de trabajo convenido. Si se le pagara periódicamente y renunciara, el empleador no tendría ningún remedio legal práctico. Los contratos de construcción suponen que el constructor ha completado el trabajo contratado pero el cumplimiento completo difiere de lo convenido en detalles menores; en tal caso, se remite al propietario a su remedio de pago de daños contra el constructor por las razones explicadas en el capítulo IV.

Horwitz sostiene a continuación que la regla aplicable en los contratos de construcción promovía el espíritu de empresa, mientras que la regla apli-

[18] Horwitz, nota 8 anterior, en 188.

cable en los contratos de trabajo no desalentaba tal espíritu porque "las disposiciones penales de los contratos de trabajo, en contraste, tienen sólo consecuencias redistributivas, ya que difícilmente podrá esperarse que disuadan a las clases trabajadoras de efectuar la venta de sus servicios en una economía de subsistencia".[19] Los Estados Unidos no sólo no tenían en el siglo XIX una economía de subsistencia, sino que no se podía inducir a quienes vivían al nivel de subsistencia a firmar contratos en cuyos términos aceptaran no recibir salarios durante todo un año. Y, si los trabajadores de esa época estuvieran viviendo al nivel de subsistencia, ninguna forma de contrato podría redistribuir la riqueza de ellos a los empleadores, pues no tendrían riqueza alguna.

Se ha sugerido que el surgimiento de la doctrina de la imposibilidad y otras relacionadas en el derecho contractual demuestra la declinación de los principios de mercado libre del derecho desde el siglo XIX.[20] Esto no es correcto. Como hemos visto, la excusa de la imposibilidad y otras relacionadas son necesarias para volver eficiente al derecho contractual; están implícitas en la teoría económica positiva del derecho contractual. Este error es especialmente desconcertante en vista de la afirmación del mismo autor en el sentido de que el derecho contractual del siglo XIX favorecía consistentemente a la parte ejecutante,[21] que, tal como sería de esperarse, estaría a favor de la doctrina liberal de las excusas.

VIII.4. El derecho común como costumbre: el cuestionamiento de Hayek del análisis económico del derecho común

El análisis económico del derecho común fue cuestionado por un famoso economista —Friedrich Hayek— cuya teoría[22] se puede resumir de la siguiente manera. Existen dos formas de establecer normas para guiar la conducta humana. En una, a la que Hayek llama "racionalismo constructivista", éstas se prescriben de forma jerárquica por una legislatura, una burocracia o un poder judicial: en otras palabras, por expertos que reúnen la información necesaria para formular de manera analítica el mejor conjunto de normas. Hayek rechaza este enfoque, del cual el análisis económico del derecho común es una versión, porque requiere demasiada información. El otro enfoque es el de la costumbre, o las normas sociales (véase después VIII.8), y es preferible, según Hayek, porque es un "orden espontáneo", a diferencia de uno generado mediante un plan o diseño. Los mercados son otro ejemplo de "orden

[19] *Idem*.

[20] Grant Gilmore, *The Death of Contract,* 80-82, 94-96 (1974).

[21] Nota 9 anterior.

[22] Véase Friedrich A. Hayek, "Law, Legislation and Liberty: A New Statement of the Liberal Principles of Justice and Political Economy", vol. 1: *Rules and Order* (1973). En Bruno Leoni, *Freedom and the Law* (3ª ed., 1991), se propuso una teoría similar del derecho.

espontáneo" en el sentido de Hayek. Éstos surgieron hace miles de años; no se inventaron o diseñaron; y su operación no implica una planificación centralizada.

La aversión tradicional de los jueces a dar apariencia de ser creativos llevó a los jueces del derecho común inglés a decir que no estaban haciendo reglas nuevas del derecho, sino que simplemente hacían cumplir la costumbre inmemorial. Hayek interpretó esta afirmación de forma literal y argumentó que lo único que un juez tiene derecho a decidir es "si la conducta objeto de controversia concuerda con […] la costumbre establecida de la que [las partes] debieran haber tenido conocimiento".[23] De forma alternativa, pero equivalente, el deber de los jueces es hacer cumplir las expectativas creadas por la costumbre. Hayek reconoce que existen lagunas en las reglas legales y, lo que es lo mismo, que "surgirán constantemente nuevas situaciones en las que las reglas establecidas son inadecuadas", lo que requiere de la "formulación de reglas nuevas" de parte de los jueces. Sin embargo, éstos deben llenar estos vacíos con la costumbre. No deben involucrarse en "mantener el equilibrio de los intereses particulares influenciados [por las reglas] en función de su importancia" o preocuparse por "los efectos de la utilidad de [las reglas] en casos particulares". De hecho, "ni los jueces ni las partes involucradas necesitan saber algo acerca de la naturaleza del orden general resultante, o acerca de cualquier 'interés de la sociedad' a la que sirven".[24]

El derecho contractual es el paradigma del derecho como costumbre. Sólo provee un marco de acción privada, dejando la identidad de las partes, así como el precio y demás condiciones del contrato, a la determinación privada. Se abstrae de todos los detalles de las interacciones voluntarias de las personas y así maximiza su libertad, su "espontaneidad".

Pero, al insistir en la obediencia a la costumbre, Hayek pasa por alto el hecho de que las costumbres suelen apoyar las actividades cooperativas que perjudican a la sociedad en su conjunto. En el capítulo VI vimos que el nivel habitual de cuidado en una industria podría ser subóptimo. Abandonadas a sus propios medios, las empresas competidoras podrían desarrollar una costumbre de condenar una reducción de precios como poco ética. Los fabricantes, si se les permitiera, desarrollarían la costumbre de ignorar la contaminación que provocaron (a menos, quizá, que afectara a sus trabajadores y a sus clientes). Y, al ser acéfalas (no existe un "dotante de costumbres" análogo a la asamblea legislativa, que es la dotante de la ley), las costumbres tienden a cambiar muy lentamente. De modo que, si las prácticas económicas u otras prácticas sociales están cambiando rápidamente, la costumbre podría dejar de estar al día y volverse un obstáculo para el progreso.

[23] Hayek, nota 22, en 87.

[24] Estas citas son de *ibid.* En 115-121 *passim*.

VIII.5. El derecho primitivo

Una de las razones de que la responsabilidad por negligencia en los accidentes ferroviarios ha parecido a los historiadores del derecho un ejemplo de la inclinación del derecho en favor de las empresas es que, en el siglo XVIII, antes del advenimiento del ferrocarril, la responsabilidad por los accidentes había sido en gran medida una responsabilidad estricta. El siglo XIX fue la culminación de una prolongada tendencia en contra de la responsabilidad estricta en el derecho: en el derecho contractual, en el penal, al igual que en el de los cuasidelitos. Esa tendencia tiene una explicación económica.[25] La difusión de la alfabetización o dominio de la lengua y el crecimiento del conocimiento científico han fortalecido las capacidades de los tribunales para la determinación de los hechos, reduciendo así los costos de la información acerca de los méritos de las disputas legales. Ocurrieron también grandes avances en el mercado de la provisión de seguros. Como vimos en el capítulo VI, la responsabilidad estricta requiere menos información que la responsabilidad por negligencia, porque evita lo que es a menudo una difícil investigación fáctica y analítica, y provee (en el derecho aplicable a los cuasidelitos y el de los contratos, no en el derecho penal) una forma del aseguramiento. Estas dos ventajas eran mayores en el pasado que ahora; y dado que, como sabemos, una responsabilidad estricta implica ciertos costos que la responsabilidad por negligencia no tiene, no es sorprendente que, puesto que los beneficios de la responsabilidad estricta han disminuido, se haya generado un movimiento hacia su abandono. De igual modo, el derecho primitivo entendía los contratos en términos muy literales; las doctrinas de mejora, tales como el cumplimiento sustancial, el error mutuo y el impedimento colateral por promesa tuvieron que esperar a las mejoras en el procedimiento judicial que permitieran a los jueces adaptar las obligaciones contractuales a las circunstancias reales en que se llevaban a cabo los contratos.

La dificultad que tenían los tribunales para hacer determinaciones de hecho explica su fuerte dependencia de la ayuda sobrenatural, como en el juicio medieval por prueba, donde, por ejemplo, el acusado podría ser absuelto si su brazo sanaba rápidamente después de arrojarlo al fuego. Al parecer, los

[25] Richard A. Posner, *The Economics of Justice,* 199-203 (1981). *Ibid.,* parte 2, es un estudio económico de las instituciones sociales y legales primitivas y antiguas. Algunos estudios más recientes son los de David Lock Hall y F. Douglas Raymond, "Economic Analysis of Legal Institutions: Explaining an 'Inexplicable' Rule of Roman Law", 61 *Ind. L. J.,* 401 (1986); Saul Levmore, "Variety and Uniformity in the Treatment of the Good-Faith Purchaser", 16 *J. Legal Stud.,* 43 (1987); Bruce L. Benson, "Legal Evolution in Primitive Societies", 144 *J. Institutional and Theoretical Econ*., 772 (1988); Geoffrey P. Miller, "Ritual and Recognition: A Legal-Economic Interpretation of Selected Biblical Texts", 22 *J. Legal Stud*., 477 (1993); Robert C. Ellickson y Charles Dia. Thorland, "Ancient Land Law: Mesopotamia, Egypt, Israel", 71 *Chi.-Kent L. Rev*., 321 (1995); Francesco Parisi y Giuseppe Dari-Mattiacci, "The Rise and Fall of Communal Liability in Ancient Law", 24 *Int'l Rev. Law and Econ.*, 489 (2004).

jueces (mas no los acusados) notaron que el procedimiento era incorrecto y, por lo tanto, supusieron que cualquier persona que se negara a la prueba (y de este modo se considerara culpable, así como hoy día el rehusarse a tomar una prueba de alcoholemia establece la culpabilidad de conducir bajo la influencia del alcohol o de drogas que alteran el estado mental) era culpable, y cualquier persona que accediera a tal cosa era inocente, y por consiguiente se aseguraron de que la prueba no fuera suficientemente grave para "demostrar" que el acusado era culpable.[26]

Si arrojamos el haz de la luz económica más atrás aún en el tiempo, podremos entender mejor los orígenes del derecho. Aun antes de que haya un gobierno, hay una forma reconocible de la responsabilidad estricta: la venganza. Si una persona ataca o roba a otra, es muy probable que la víctima ejerza represalias y que esta amenaza provea la disuasión contra el comportamiento agresivo. Los costos de la información en una sociedad sin instituciones de adjudicación dificultarán la diferenciación entre los daños intencionales, por descuido o negligencia, por una parte, y los accidentes inevitables a un costo razonable, por otra parte, de modo que la responsabilidad tenderá a ser estricta. La represalia es probable, aunque su costo sea mayor que cualquier ganancia que la víctima pudiera esperar obtener de la represalia contra su agresor. Esto podría parecer irracional. Pero, si prevaleciera el pensamiento "racional" —es decir, si los costos de ser una víctima se trataran como hundidos, de modo que la víctima ejercería represalias sólo si la ganancia esperada superara al costo esperado—, raras veces se ejercerían represalias y el sistema de la disuasión se derrumbaría. La disposición a ejercer represalias aunque los costos esperados superen a los beneficios esperados en el momento en que deba tomarse la decisión de ejercer o no represalias (es decir, *después* de que ha ocurrido la agresión contra la víctima) —la clase de disposición que mantuvo el "balance del terror" durante la Guerra Fría con la creación, por parte de los Estados Unidos y la Unión Soviética, de la capacidad de efectuar un segundo ataque— puede ser la estrategia que minimice el costo *ex ante*. Con todo, para que ésta sea una estrategia creíble y logre por lo tanto su objetivo de disuasión, la víctima potencial debe ser capaz de *comprometerse previamente* a ejercer represalias.[27] La emoción de la venganza, que probablemente forma parte de nuestra dotación genética, provee el necesario compromiso previo. Es difícil comprender cómo, sin una predisposición genética a la venganza, podía mantenerse el orden antes de que hubiera un gobierno.

Es muy probable que la represalia se dirija contra los miembros de la familia del agresor, al igual que contra el agresor mismo o incluso en lugar de éste. Esto no sólo minimiza el costo de la represalia al otorgar al vengador más objetivos potenciales, sino que también otorga a la familia del agresor un incentivo para impedirle que agreda en primer lugar. Éste es el principio

[26] Véase Peter T. Leeson, "Ordeals" (George Mason U., Dept. of Econ., 10 de enero de 2010).

[27] Compárese con la discusión de los precios de depredación, más adelante, en X.9.

del castigo colectivo, y sigue desempeñando un papel importante en el derecho. Además de los ejemplos de VII.3, considérese el caso de Ybarra *vs.* Spangard.[28] Al demandante se le practicó una apendicectomía bajo anestesia general y, cuando despertó, descubrió que se le había causado una lesión grave en el hombro. Demandó a los médicos y las enfermeras que habían estado presentes en la operación, pero no pudo identificar cuál de ellos había causado la lesión. El tribunal sostuvo que, puesto que era probable que uno o más de ellos hubiese causado esa lesión, les correspondía a ellos la carga de la prueba para que explicaran lo que había ocurrido; no pudiendo liberarse de la carga, ellos perderían el juicio. La decisión le daba a cada uno de ellos, excepto al culpable real, un incentivo para revelar la identidad del culpable.

El castigo colectivo no sólo consigue enlistar a más responsables del cumplimiento —las víctimas potenciales de castigos colectivos (familiares u otros asociados del delincuente inicial)— sino que también reduce los costos de detección. El responsable del cumplimiento no tiene que determinar quién es el culpable: basta con que pueda definir a algunos miembros de la familia u otro colectivo al que pertenezca el culpable. Pero esto puede diluir la disuasión marginal; es menos probable que se castigue al delincuente si la represalia de las víctimas se distribuye entre los asociados del delincuente en lugar de que se enfoque en él por completo.

Un sistema de venganza para mantener la ley y el orden tiene una debilidad más grave: sacrifica a las economías importantes facultadas por la especialización. No todas las personas son muy aptas para ser vengadores (ése es uno de los temas de *Hamlet*); tampoco es eficiente que todos los varones adultos sean vengadores de medio tiempo. Un sistema moderno de justicia penal cuenta con profesionales que buscan, enjuician y castigan a los criminales, no por odio u obligación familiar, sino porque se les paga por hacerlo. Sin embargo, estos profesionales dependen en gran medida de los particulares para que ayuden en el cumplimiento, incluyendo a las víctimas de delitos, a las familias de las víctimas, a los testigos, a los miembros del jurado y a otros que por lo general carecen de interés monetario alguno en la captura y enjuiciamiento de los delincuentes. No existe incompatibilidad entre la teoría económica del castigo y el reconocimiento de que la emoción desempeña un papel en el cumplimiento de la ley como un motivador alternativo a los incentivos de carrera para llevar a los delincuentes ante la justicia.

La siguiente etapa después de la venganza en el mantenimiento del orden en las sociedades primitivas consiste en permitir que el agresor pague a la víctima de su agresión para que no tome represalias. En otras palabras, un sistema de pagos de transferencia sustituye al sistema socialmente más costoso de la fuerza; la analogía es con la sustitución de la prisión por las multas (cuando sean viables). Pero, en lugar de variar el pago según el daño particular

[28] 154 P.2d 687 (Cal., 1944).

de la víctima particular, como en el moderno sistema de los cuasidelitos, los pagos (llamados "aranceles", pero que no deben confundirse con los impuestos cargados a las importaciones) se hacen de acuerdo con una tarifa fija: tanto por matar a un hombre, tanto por matar a una mujer o a un niño, tanto por tumbar un diente, etc. Dicho sistema representa un ajuste razonable a los altos costos de información, sobre todo porque, cuando los victimarios potenciales no pueden calcular la medida precisa de la lesión por adelantado, los recursos gastados en la determinación de esa medida después del hecho no alteran los incentivos de los victimarios.[29]

La economía puede iluminar curiosidades históricas tales como el sistema sorprendentemente estable del derecho en la Islandia medieval, una sociedad que carecía virtualmente de Estado. La adjudicación por jueces profanos (equivalentes a nuestros jurados) floreció junto con el arbitraje informal, aun cuando ni el cumplimiento de los fallos judiciales ni el de los arbitrales corría a cargo de funcionarios públicos: no había ninguno. Si un litigante perdedor se negaba a obedecer un fallo (los fallos judiciales más comunes eran la proscripción y el exilio), el ganador tendría que reunir a sus parientes para hacerlo cumplir. Pero eso no volvía inútil el recurso del litigio, porque un fallo tendría a menudo suficiente fuerza de persuasión para desalentar a los parientes del perdedor a fin de que no lo protegieran contra los parientes del ganador, y para alentar a estos últimos a ayudar al ganador para el cumplimiento del fallo. Un factor facilitador era que los islandeses reconocían el parentesco a través del lado masculino y el femenino de la familia. Por lo tanto, todos tenían muchos parientes y esto hacía más probable que los litigantes tuvieran algunos parientes en común, quienes ayudarían a bajar el calor de la disputa, promoviendo el arreglo o permitiendo que la ley siguiera su curso.[30]

Antes de que existiera el seguro de mercados o la seguridad social, o incluso los contratos que proveían un tipo de seguro al exigir a la parte que realizaba el contrato que compensara a la otra parte en caso de incumplimiento, aun cuando el incumplimiento estuviera más allá de la capacidad de la parte ejecutante de haberlo evitado, existía un seguro informal que las familias, y otros grupos pequeños, proporcionaban a sus integrantes. A través de un sistema de intercambio recíproco, los parientes y los miembros del clan se protegían entre sí del hambre y de otras carencias graves comunes en las sociedades primitivas. La resistencia milenaria a la "usura" (en el sentido original de cobrar intereses por un préstamo, a diferencia de su sentido moderno de interés excesivo) puede reflejar la incongruencia de cobrar intereses por otor-

[29] Louis Kaplow y Steven Shavell, "Accuracy in the Assessment of Damages", 39 *J. Law & Econ.*, 191 (1996). El método del "arancel" es el que se usa en la compensación de los trabajadores en lugar de la evaluación individualizada de los daños que se emplea en los casos de los cuasidelitos convencionales.

[30] Véanse David Friedman, "Private Creation and Enforcement of Law: A Historical Case", 8 *J. Legal Stud.*, 399 (1979); Thráinn Eggertsson, *Economic Behavior and Institutions*, 305-310 (1990); Richard A. Posner, *Overcoming Law*, cap. 14 (1995).

gar la ayuda con la que el prestamista espera ser correspondido algún día cuando se encuentre en problemas. Por otra parte, en tanto que había una norma sólida de intercambio recíproco, podía no confiarse en que los prestatarios retribuyeran sus préstamos, porque sabrían que incluso si derrocharan los recursos del préstamo y terminaran sin un centavo todavía podrían contar con la ayuda de sus parientes o del clan.[31]

VIII.6. El derecho y la modernización

Hay una conciencia creciente de que el hecho de que los gobiernos de los países pobres no provean la infraestructura legal básica de una economía capitalista es un factor importante para que estos países sigan siendo pobres. La infraestructura requerida consiste fundamentalmente en una maquinaria para el cumplimiento de los derechos legales, especialmente los derechos de propiedad y de contrato. En su forma ideal, la maquinaria consiste en jueces profesionales competentes, éticos y bien pagados que administren reglas bien diseñadas para la promoción de la actividad comercial, que sean suficientemente numerosos para decidir los casos sin una demora interminable, que estén aislados de la interferencia de los poderes legislativo y ejecutivo, cuyos fallos se hagan cumplir confiablemente por actuarios, alguaciles, policías u otros funcionarios (de nuevo, competentes, éticos y bien pagados), y que estén asesorados por abogados competentes, éticos y bien pagados, que operen con un trasfondo de reglas y prácticas tales como estándares de contabilidad, oficinas de estadísticas demográficas y registros públicos de títulos de la tierra y de intereses de seguridad, que les permitan resolver las disputas de hechos con razonable corrección y a un costo razonable para los litigantes.

No obstante, incluso grandes desviaciones de este ideal podrían no impedir gravemente la eficiencia económica. Esto lo sugiere el éxito económico de estados de la Unión Americana que tienen jueces elegidos y politizados con cuestionable competencia profesional; el éxito económico de naciones asiáticas tales como China, donde el imperio de la ley es débil (en China, apenas puede decirse que existe); el hecho de que Inglaterra, que tiene uno de los mejores sistemas judiciales del mundo, fuese durante muchos decenios uno de los países de menor crecimiento económico entre los desarrollados; y el hecho de que la India, cuyo ordenamiento jurídico sigue el modelo inglés y que tiene una vasta población de abogados, haya crecido menos que China, un país que inició su despegue económico en el decenio de 1980 en una época en que sólo contaba con un ordenamiento jurídico rudimentario.

¿Cómo explicar estas anomalías? Una posible explicación es que, en virtud de que el régimen jurídico "ideal" es una empresa cara para un país pobre, el

[31] *Cf.* Jared Rubin, "Social Insurance, Commitment, and the Origin of Law: Interest Banks in Early Christianity", 52 *J. Law & Econ.*, 761 (2009).

intento de obtenerlo habría significado elevados costos de oportunidad sociales. Otra es que un régimen jurídico cumple muchas más funciones que la de simplemente hacer respetar los derechos contractuales y de propiedad, incluido el respeto forzado de leyes malas que reducen la eficiencia económica.[32] Además, hay muchos sustitutos privados para la protección legal de los derechos de propiedad y de contratos. Se incluye aquí el arbitraje, con o sin el respaldo legal del fallo del árbitro; la reputación, que a su vez permite las represalias, incluidas por ejemplo el registro en listas negras de quienes incumplan sus contratos y otras formas de ostracismo; la fusión entre empresas, de modo que las disputas manejadas anteriormente por el ordenamiento jurídico se vuelven disputas puramente internas que resuelven los administradores; el monopolio bilateral, que puede proveer un sustituto para los contratos de trabajo legalmente exigibles (véase después XII.4); las tácticas de brazo fuerte del tipo que se encuentra en los mercados ilegales, tal como el narcotráfico ilegal en los Estados Unidos; y el altruismo, que permite la operación eficaz de muchas empresas de propiedad familiar fuera de un marco legal,[33] sobre todo si hallan a sus trabajadores y clientes en una comunidad étnica o religiosa muy unida. El altruismo y el ostracismo van de la mano, ya que juntos crean una comunidad de lazos estrechos lo bastante compacta para ser manejable, unida contra el mundo: "We few, we happy few, we band of brothers" (*Henry V*).*

La importancia de los sustitutos privados del derecho se revela por el hecho de que la propiedad y el contrato son métodos de coordinación y optimación de la actividad económica que anteceden ampliamente a las instituciones jurídicas estatales y formales. Incluso en la cultura altamente litigiosa de los Estados Unidos, la gran mayoría de las disputas de contratos y propiedades se resuelven informalmente y ello ocurriría aun sin un marco de amenaza de recurrir a la ley, porque esa amenaza es a menudo inocua en virtud de la carestía de los procedimientos legales.

Sin embargo, *cierta* aproximación al orden jurídico "ideal" contribuye sin duda considerablemente al crecimiento económico.[34] Cuando la ley es débil o inexistente, el respeto forzoso de los derechos de propiedad y de contratos

[32] Esto podría explicar la correlación negativa encontrada en Kevin M. Murphy, Andrei Shleifer y Robert W. Visny, "The Allocation of Talent: Implications for Growth", 106 *Q. J. Econ*., 503 (1991), entre el número de abogados de una nación y la tasa de crecimiento económico de esa nación. La correlación es potencialmente engañosa porque gran parte de la producción de abogados consiste en bienes fuera del mercado; pero éstos podrían no ser tan importante en los países pobres como en los ricos.

[33] Véase Paul H. Rubin, "Growing a Legal System in the Post-Communist Economies", 27 *Cornell Int'l L. J*., 1, 17-25 (1994).

* "Nosotros pocos, felices pocos, nosotros, grupo de hermanos": William Shakespeare, *Enrique V*, acto IV, escena III, trad. de José María Valverde, Planeta, Barcelona, 1988.

[34] Véanse, por ejemplo, Robert J. Barro, "Economic Growth in a Cross Section of Countries", 106 *Q. J. Econ*., 407, 437 (1991); Gerald W. Scully, "The Institutional Framework and Economic Development", 96 *J. Pol. Econ.,* 652 (1988); Daniel M. Klerman y Paul G. Mahoney, "The Value

depende con frecuencia de la amenaza de la violencia y a veces de su aplicación efectiva, de los métodos enredados de autoprotección y de las alianzas familiares y de clanes que podrían ser económicamente disfuncionales e interferir también con la justicia legal impersonal, la cual tiene un valor económico independiente que se discute después en VIII.9. Estos sustitutos generan también un sesgo en contra de las empresas nuevas, que no tienen una reputación que puedan emplear para persuadir a los socios contractuales potenciales de su confiabilidad, por lo general un prejuicio contra los foráneos (ya que el derecho es un sustituto de la confianza), así como un sesgo a favor de los intercambios simples, simultáneos, sobre las transacciones más complejas, porque, como vimos en el capítulo IV, entre más se aproxime a la simultaneidad el cumplimiento de las partes de un contrato, menor será la necesidad de un remedio legal contra el incumplimiento.

A menos que haya un mecanismo eficaz para el cumplimiento imparcial de los contratos, las personas estarán renuentes a hacer negocios con extraños. En una cultura como ésta, el trueque y otras formas de negociaciones recíprocas basadas en (y constructoras de) relaciones de confianza, junto con los lazos del clan y las relaciones de la familia extensa, organizarán una actividad económica importante y harán que sea natural que las personas consideren a los funcionarios públicos como personas que venden servicios públicos a sus amigos y parientes. En cambio, la corrupción debiera ser poco común en un sistema de libre mercado con tribunales que hacen cumplir los contratos de manera honesta y confiable.

Un país pobre podría no ser capaz de brindar un régimen jurídico bueno, pero sin tal sistema es posible que nunca se vuelva suficientemente rico para proveerlo. Por lo tanto, puede afirmarse que, en vista de que es mucho más costosa y consumidora de tiempo la creación de instituciones jurídicas eficientes que la promulgación de reglas eficientes para la administración de las instituciones ineficientes existentes, la reforma debiera concentrarse en esta última.[35] La creación y difusión de una regla implica costos fijos pequeños y (como otros bienes de información) costos marginales insignificantes, mientras que las instituciones jurídicas requieren grandes insumos de mano de obra educada y cara. Esto implica a su vez que la estrategia de primero las reglas es mejor entre más populoso sea un país, porque el costo de la creación de una regla es invariable ante el número de personas a quienes se aplicará la regla.[36] China, el país más poblado del mundo, ha seguido la estrategia de las reglas primero.

of Judicial Independence: Evidence From Eighteenth Century England", 7 *Am. Law and Econ. Rev.*, 1 (2005).

[35] Véase Jonathan R. Hay, Andrei Shleifer y Robert W. Vishny, "Privatization in Transition Economies: Toward a Theory of Legal Reform", 40 *European Econ. Rev*., 559 (1996).

[36] El punto general es que, dado que la regulación legal de cualquier tipo implica costos fijos, cuanto mayor sea la jurisdicción regulatoria será más barata la regulación debido a que los cos-

Debemos distinguir dos tipos de eficiencia de las reglas. Una regla es *sustantivamente* eficiente si establece un precepto que, de ser obedecido, interiorizará una externalidad o promoverá de otro modo la asignación eficiente de los recursos; una regla que prohíba el uso de la propiedad de otra persona sin su consentimiento (es decir, que prohíba el allanamiento) es un ejemplo. Una regla es *procedimentalmente* eficiente si está diseñada para reducir el costo o aumentar la corrección del uso del régimen jurídico. Tal es el caso, por ejemplo, de una regla que requiera que los contratos se celebren por escrito para que sean legalmente exigibles, una regla de que ninguna reclamación de infracción de derechos es procedente si se presenta después de tres años de cometida la infracción, una regla que requiera que ciertas disputas, digamos entre empleadores y empleados o entre corredores de valores y sus clientes, deberá presentarse al arbitraje vinculante, y una regla que faculte al ganador de un juicio de daños a recibir intereses sobre el fallo a la tasa del mercado desde la fecha de presentación de la demanda. Las dos primeras reglas procedimentales de la lista están diseñadas para reducir los costos de información del sistema jurídico, la tercera para reducir la carga del trabajo judicial derivando ciertas disputas a un método alternativo de resolución de las disputas y la última para permitir que los jueces usen las colas para hacer frente a una gran carga de trabajo sin devaluar la compensación que el sistema legal confiere a las personas cuyos derechos han sido infringidos. El requisito del arbitraje de ciertas disputas tiene una importancia adicional como un método para alentar la formación de asociaciones comerciales y otros grupos empresariales que en una sociedad comercial son instituciones intermedias importantes entre el individuo y la familia, por una parte, y el Estado, por la otra: se mantienen entre el nepotismo y el estatismo.

La sencillez relativa con que pueden administrarse las reglas en comparación con los estándares[37] tiene dos consecuencias para la clase de poder judicial débil que es común en un país pobre. La aplicación de las reglas impone menos demandas al tiempo y la competencia de los jueces, de modo que resulta más barata y tiene mayores probabilidades de ser correcta. Las reglas facilitan también el monitoreo de los jueces y, así, reducen la probabilidad del soborno y de la influencia de la política en el proceso judicial. Entre menos discreción tenga un juez en la toma de decisiones, más fácil será que los observadores determinen si ha fallado un caso contra la ley o está favoreciendo sistemáticamente a una clase o un grupo de litigantes contra otro.

Otras medidas económicas para la disminución de la corrupción y la incompetencia judicial incluyen:

tos fijos se pueden extender más. Véase, por ejemplo, Casey B. Mulligan y Andrei Shleifer, "The Extent of the Market and the Supply of Regulation", 120 *Q. J. Econ.*, 1445 (2005).

[37] Véase después, en XXI.3, una comparación económica más completa entre las reglas y los estándares como métodos de la gobernanza legal.

1. *Apoyar la compensación judicial* mediante la creación de generosos derechos de pensión que se pierden si el juez es despedido del cargo por incompetencia o corrupción. Esto aumenta el castigo del juez por su mal comportamiento.[38]
2. *Hacer que los jueces se integren en colegios*, en lugar de que obren solos, o haciéndolos sentarse con jurados. En ambos casos, aumentarán los costos de transacción del soborno y la probabilidad del descubrimiento.
3. *Disminuir los castigos legales* a fin de reducir la ganancia del soborno de los jueces y otros funcionarios para que no los impongan. Si la multa por exceder los límites de velocidad fuese de 1 000 dólares, sería grande el incentivo de un policía para pedir, y del infractor para pagar, un soborno a fin de evitar una boleta de infracción. Ésta es otra razón para que el sistema de la justicia penal no combine el castigo más severo posible con la probabilidad más baja de imponerlo (véase el capítulo VII).
4. *Establecer prioridades*. Si el mayor peligro para la prosperidad de una nación en vías de desarrollo es la amenaza de que el gobierno confisque la propiedad de los particulares, podría tener sentido la creación de un tribunal especial cuya única misión sea ejercer funciones de contraloría del gobierno (el enfoque francés; el *Conseil d'État* es ese tribunal). Los jueces de *este* tribunal tendrán que ser competentes, éticos y bien pagados, pero, dado que la jurisdicción del tribunal estará tan circunscrita, serán moderados los recursos totales de capital humano y físico requeridos para el equipamiento del tribunal. Y, si el tribunal se limita a las cuestiones puramente económicas, las autoridades políticas podrían estar dispuestas a tolerar su independencia, sobre todo si entienden cuánto puede hacer tal institución por la prosperidad del país.
5. *Establecer la finalidad*. En muchos países los litigios se ven afectados por retrasos severos, en parte debido a la falta de apreciación de la importancia del establecimiento de la finalidad —de llevar los litigios a un cierre oportuno y definitivo—. Un ejemplo de ello es Brasil, que tiene un problema de retraso judicial de proporciones increíbles desde el punto de vista de los Estados Unidos. En 2010, más de 70% de los casos pendientes en los tribunales brasileños se habían presentado en años anteriores; a menudo muchos años antes.[39] Desde 1988, "la acumulación de casos pendientes en la agenda judicial se ha incrementado diez veces y, en consecuencia, los retrasos en los juicios se han duplicado con creces […] Algunos casos están pendientes desde la década de 1940 […] Las estadísticas conservadoras estiman el número de deman-

[38] Véase Gary S. Becker y George J. Stigler, "Law Enforcement, Malfeasement, and Compensation of Enforcers", 3 *J. Legal Stud*., 1, 6-13 (1974).

[39] Véase CNJ (Conselho Nacional de Justiça), *Justice in Numbers*, 2010, 6 (2010).

> das en espera del fallo final en más de 50 millones. Entre 1995 y 1999, se presentaron 32.2 millones de procesos [casos] en los tribunales brasileños. Sin embargo, sólo 22.6 millones de éstos se decidieron durante el mismo periodo. Esto deja un déficit de casi 10 millones de procesos sin juzgar."[40] Son varias las razones de este retraso, pero una de ellas es que el sistema judicial brasileño sólo tiene un concepto débil de finalidad. Además de permitir las apelaciones interlocutorias de manera indiscriminada, la cultura jurídica de Brasil también autoriza la reapertura en pareja medida. Los costos en sobrecarga judicial son enormes.

La corrupción, sin limitarla a la corrupción judicial, florece cuando una economía está muy regulada pero el marco legal es débil. Entre más regulada esté la economía, las restricciones serán más fastidiosas, y esto creará una demanda de formas para evitar cumplir con las mismas. Sobornar a quienes se encargan de hacer cumplir las restricciones es una de esas formas. Cuanto más débil sea el marco legal, será más difícil para el gobierno evitar el soborno, un delito "sin víctimas" clásico porque se trata de una transacción voluntaria. Se requiere de un mecanismo legal sofisticado para detectar y castigar estos delitos.

¿Cuán destructiva es la corrupción, económicamente? (y, por lo tanto, ¿qué tan importante es un régimen jurídico eficiente?). Quizá mucho, debido a que la mayor carga recae sobre la innovación.[41] Las actividades innovadoras son frágiles y arriesgadas. El soborno que exigen los funcionarios públicos podría matar a la gallina de los huevos de oro desde el comienzo. Además, si la empresa innovadora tiene éxito, es probable que los funcionarios se apoderen de una gran parte de los beneficios, mientras que, si falla, los costos caerán por completo sobre todos los partícipes. Estas perspectivas vuelven poco atractiva a la innovación.

VIII.7. El imperio de la ley

Las ideas fundamentales de la sección anterior se resumen en el concepto del "imperio de la ley" (la "racionalidad formal" del derecho en la útil frase de Max Weber). El filósofo John Rawls identificó cuatro elementos del concepto: el derecho *1)* debe ser susceptible de cumplimiento por parte de aquellos a quienes se dirige; *2)* debe tratar por igual a las personas que se encuentran en una situación similar con respecto a una norma jurídica determinada; *3)* debe ser público, y *4)* debe fijar un procedimiento por el que se establezca la verdad de cualesquiera hechos que sean necesarios para la aplicación de la norma

[40] Augusto Zimmermann, "How Brazilian Judges Undermine the Rule of Law: A Critical Appraisal", 11 *Int'l Trade and Bus. L. Rev.*, 179, 192 (2008).

[41] Véase Kevin M. Murphy, Andrei Shleifer y Robert W. Vishny, "Why Is Rent-Seeking So Costly to Growth?", 83 *Am. Econ. Rev. Papers and Proc.*, 409, 412-413 (mayo de 1993).

jurídica.[42] Curiosamente, lo que falta es el origen del derecho: si el derecho de una sociedad es el capricho de un dictador, ¿son, aun así, suficientes los cuatro elementos de Rawls para establecer que la sociedad se encuentra bajo el imperio de la ley, en caso de que todos estén presentes? Una idea más tradicional del imperio de la ley, con base en la distinción aristotélica entre la justicia correctiva y la distributiva, sólo requiere de la presencia del segundo de los cuatro factores de Rawls. Véase antes I.2. Más adelante abundaremos sobre el tema.

Los elementos de Rawls resultan ser centrales para el concepto económico del derecho. La función primordial del derecho, en una perspectiva económica, es la de modificar los incentivos. Esto implica que el derecho no ordena lo imposible, porque una orden imposible de cumplir no modificará el comportamiento.

El requerimiento de que el derecho debe tratar igualmente a los iguales es otra manera de decir que el derecho debe tener una estructura racional, porque es irracional tratar diferentemente a cosas que son lo mismo. La teoría económica es un sistema de lógica deductiva: cuando se la aplica correctamente produce resultados que son congruentes entre sí. En la medida en que el derecho tiene una estructura económica implícita, debe ser racional; debe tratar de igual manera a los casos semejantes.

Pero hay algo más en la noción de la "igualdad frente al derecho", y ese "más" liga el concepto del imperio de la ley al de la protección igual de las leyes. Entre más general es la ley, más difícil resulta para el gobierno el castigo de los disidentes, la explotación de individuos impopulares o la comisión de otros abusos. Supongamos que el gobierno quiere dañar a A, pero se ve constreñido por el concepto de la igualdad ante la ley a hacerlo por medio de una reglamentación que también dañará a B, porque B cae dentro de la clasificación hecha por la regla junto con A. Si el gobierno no desea dañar a B, el requerimiento de la generalidad protegerá a A al aumentar el costo, para el gobierno, de la opresión que quiere imponerle.

La ley también debe ser pública. Si el contenido de una ley sólo se conociera después de que ocurrieran los sucesos a los que puede aplicarse, la existencia de la ley no tendría ningún efecto sobre la conducta de las partes que estuvieran sujetas a ella. La amenaza que no es comunicada no puede disuadir. La principal excepción es la teoría de la prevención o incapacidad que subyace en algunas teorías del derecho penal y que, como vimos en el capítulo anterior, tiene una justificación económica.

La teoría económica del derecho presupone una maquinaria para determinar la existencia de los hechos necesarios para la aplicación correcta de una ley. El efecto disuasivo de la ley se debilita (y en el límite desaparecería) si se hace cumplir sin tomar en cuenta si están presentes las circunstancias a

[42] John Rawls, *A Theory of Justice*, 237-239 (1971).

las que trataba de aplicarse la ley. Supongamos que hay una ley contra quienes fijan los precios, pero no se hace ningún esfuerzo para determinar quién está haciéndolo; en cambio, se selecciona al azar a 10 000 personas y se les castiga por fijar precios. Entonces no habrá ninguna desistencia del hecho. La única diferencia entre el fijador de precios y la persona que no los fija es que el primero obtiene ganancias por ello; la responsabilidad esperada de las dos es la misma.

La idea del imperio de la ley, cuyos elementos ya se han esbozado, se origina en el concepto aristotélico de la "justicia correctiva", un concepto en sí mismo rico en importancia económica. Aristóteles sostenía que, si alguien se involucra en una conducta injusta que perturba el equilibrio preexistente de la riqueza o de otras ventajas entre esta y otra persona, en perjuicio de esta última, la otra persona tiene derecho a algún tipo de compensación que, en la medida de lo posible, restablecerá el equilibrio preexistente o, en otras palabras, rectificará la desviación del equilibrio causada por el acto injusto.

El concepto aristotélico de la justicia correctiva es muy abstracto. No se especifica lo que se considerará como una conducta injusta; ni las formas de compensación que se considerarán apropiadas. La sustancia está en un corolario que Aristóteles elaboró a partir del concepto. El corolario es que la justicia correctiva no considera las cualidades de la persona, el mérito o abandono del malhechor y de su víctima. La víctima podría ser un hombre malo y el malhechor un hombre bueno, si tomamos en consideración el carácter de una persona y toda su trayectoria: la suma de todas sus buenas y malas acciones y no sólo el episodio concreto que produjo el daño a la víctima. Sin embargo, la víctima tiene derecho a una compensación. La razón de que esto sea un corolario de la justicia correctiva antes que un principio independiente de la justicia es que la justicia correctiva busca restablecer el equilibrio preexistente en lugar de cambiarlo. El tribunal no utiliza la ocasión para enriquecer o empobrecer al malhechor o a la víctima con base en un juicio acerca de sus méritos o abandonos, al margen de las circunstancias del propio daño. Pues esto no restablecería el equilibrio preexistente para las partes; crearía un equilibrio nuevo.

Este concepto de justicia "que no toma en cuenta a las personas", un concepto simbolizado de manera célebre por la estatua de la justicia como una diosa con los ojos vendados, es indispensable para la eficiencia. Si el obtener la compensación de los perjuicios dependiera de nuestra reputación, la gente invertiría recursos excesivos en ser popular, muy apreciada e influyente. En la medida en que tales inversiones tomen la forma de la realización de cosas realmente buenas, mejorarían el bienestar social. Sin embargo, a menudo sería más fácil obtener una buena reputación tratando de congraciarse con los poderosos, aliándose con ellos a través del matrimonio, evitando posturas impopulares y tomando otras medidas no relacionadas, con frecuencia perjudiciales incluso, para bien de la sociedad. La consecuencia sería una socie-

dad infestada por el clientelismo, el exclusivismo, el nepotismo y el soborno. Incluso cuando una persona obtuviera una buena reputación por los medios adecuados, una vez que tuviera esa reputación podría utilizarla para causar perjuicios con impunidad a las personas que carecieran de ella. Quienes no tuvieran amigos serían una clase de bandidos de quienes cualquiera de los "buenos" podría aprovecharse impunemente. Los esfuerzos serían desviados de actividades socialmente constructivas hacia la búsqueda de rentas y el establecimiento de alianzas.

VIII.8. El derecho y las normas sociales

El régimen jurídico de la Islandia medieval (véase antes VIII.5) pone de relieve la importante cuestión de la relación existente entre el derecho y las normas sociales, es decir, las reglas del comportamiento cuyo cumplimiento no es impuesto por los tribunales u otras autoridades públicas. Las costumbres, discutidas anteriormente en relación con la crítica de Hayek del racionalismo jurídico, son ejemplos de normas sociales. Sin embargo, los Estados que no tienen un gobierno o tienen un gobierno débil (la ausencia de un gobierno mundial hace del mundo moderno un buen ejemplo de una sociedad "sin Estado"), Estados que luchan para modernizarse con la desventaja de un sistema legal disfuncional, son raras veces anárquicos. Esto sugiere que definir al derecho en términos de las órdenes de un soberano u otras promulgaciones oficiales, y dejar allí el asunto, es inadecuado. Las normas sociales que son como leyes se obedecen a menudo a pesar de la ausencia de sanciones legales por su violación. A veces se cumplen por emoción, lo que es un punto vital para entender el *surgimiento* del derecho, como hemos visto.

Las normas sociales se hacen cumplir también mediante expresiones de reprobación, por el ridículo y por el ostracismo. La eficacia de las sanciones más moderadas reside en su amenaza implícita de ostracismo, es decir, de la negación de transacciones ventajosas. Una persona que viste estrafalariamente o escupe en público podría verse rechazada. Sufrirá la misma clase de pérdida que un vendedor que pierde clientes porque se descubre que los ha estado engañando, o el hablante de un idioma extranjero que no puede hacerse entender.

En virtud de que quienes aplican el ostracismo hacen cumplir las normas sin que se les pague, hay un potencial problema de "gorroneo". Sin embargo, el costo es a menudo menor o incluso negativo para los aplicadores. Cuando la violación de una norma social señala la inconveniencia de negociar con el infractor, como en el caso de quien incumple una promesa o del individuo que actúa extrañamente, es probable que el ostracismo se justifique por el costo privado y por el costo social porque evita los costos futuros del trato con el infractor de la norma. Había mucho arbitraje aun antes de que los tribunales

ejecutaran los laudos arbitrales (la Islandia medieval es un ejemplo), porque una parte en un acuerdo de arbitraje que se negara a aceptar el resultado de un arbitraje al que se había comprometido se arriesgaría al ostracismo impuesto por su comunidad comercial.[43] E, incluso si el riesgo del ostracismo indujera generalmente a las partes a cumplir sus contratos y de este modo evitara la necesidad de arbitraje, a menudo necesitarían de un árbitro para determinar si el contrato se *ha* incumplido.

Las normas sociales son más eficaces cuanto menor sea el grupo sujeto a ellas, porque las transacciones repetitivas, que no sólo reducen el costo del ostracismo para quienes lo imponen, sino que también facilitan identificar los violadores de las normas, son más frecuentes entre los miembros de un grupo pequeño. Las normas son también más eficaces entre mayor sea el costo del ostracismo y más lento sea el ritmo del cambio social. El problema del "gorroneo" que deriva de la ausencia de una autoridad centralizada que cree o cambie las normas vuelve muy difíciles los grandes o frecuentes cambios de las normas. La evolución, que es un proceso de pequeños cambios incrementales, evita este problema; y cuando el costo de una innovación normativa es pequeño, el hecho de que el innovador no pueda capturar la mayor parte de los beneficios no tiene que ser un gran obstáculo. Pero la evolución tarda mucho tiempo en generar un cambio grande, de modo que si la sociedad está cambiando rápidamente es posible que la gobernanza mediante normas no satisfaga las necesidades de la sociedad.[44] Es probable que se cumplan las tres condiciones (el grupo pequeño, el alto costo de ser condenado al ostracismo y la sociedad estática) en comunidades primitivas, aisladas. Esto explicaría la manera en que la ley pudo evolucionar como un sistema de normas sociales antes de que existiera una autoridad centralizada.

Las normas sociales que se internalizan se obedecen por un sentido de culpa o de vergüenza. Los padres motivados por el altruismo o por el estrecho interés propio, así como los maestros pagados por los padres o por el Estado, tratan de meter las normas en la cabeza de un niño, las convierten en hábitos cuya violación hará que el niño se sienta incómodo aun después de crecer hasta volverse adulto. Los incentivos para infundir las normas en los hijos y otros miembros de la familia son especialmente poderosos en las sociedades donde el castigo es colectivo, lo que es una característica común de los sistemas normativos basados en la venganza.

[43] Véase Bruce L. Benson, "An Exploration of the Impact of Modern Arbitration Statutes on the Development of Arbitration in the United States", 11 *J. Law, Econ. & Organization,* 479 (1995).

[44] En virtud de que los cambios pequeños se incorporan fácilmente al lenguaje, éste puede cambiar y en efecto cambia sustancialmente, aunque en una forma gradual. Por supuesto, debe haber *algún* beneficio positivo para que el empresario de las normas se sienta motivado para innovar; en el caso de las innovaciones del lenguaje, el beneficio consiste en permitir que el escritor exprese el significado con mayor claridad o vividez. En contraste con los cambios que ocurren *dentro* de un lenguaje, imaginemos que se trata de cambiar a una nación de una lengua a otra. Los costos serían enormes, y los beneficios muy difusos, aunque podrían exceder a los costos.

El costo del ostracismo para quien lo sufre se relaciona negativamente con el nivel del ingreso. En una sociedad rica el individuo depende menos de la buena voluntad de su comunidad particular, ya sea porque él mismo es rico o porque tiene debajo una red social de seguridad. Por lo tanto, podría ser poco realista pensar que en una sociedad rica puede reducirse la cantidad de la regulación legal sustituyéndola por la regulación normativa. Otra razón para esta conclusión es que la privacidad, un bien superior como la educación, reduce los costos privados de la violación de las normas haciendo más difícil que los ciudadanos privados, potenciales aplicadores de las normas, observen el comportamiento de otras personas. Entre más proteja la ley a la privacidad, más aumentará la demanda de la ley como un sustituto de la regulación por las normas. Un rompimiento de la regulación por las normas podría explicar el hecho de que las legislaturas consideraran necesario empoderar a los tribunales para que hicieran cumplir los fallos arbitrales.

Es de esperarse que la ley asigne una sanción a una violación de una norma buena cuando los beneficios privados de la violación son grandes o los costos privados (ya sea en términos de la pérdida de transacciones ventajosas o de la culpa) son pequeños, y cuando la violación inflige sustanciales costos sociales. A menos que se satisfagan ambas condiciones, es improbable que los beneficios de la sanción legal excedan a los costos de la regulación legal. Es por ello que la ley penaliza el homicidio, pero no a la rudeza, y los daños negligentes, pero no (en general) a los daños resultantes de un accidente puro, aunque la gente sienta culpa cuando inflige aun un daño inevitable, lo que revela que hay una norma, aunque no en general una ley, de responsabilidad estricta por infligir daños.

VIII.9. El contenido moral del derecho común

La teoría de que el derecho común se entiende mejor como un sistema para la promoción de la eficiencia económica parecerá incompleta a muchos lectores, si no es que gravemente empobrecida, sobre todo en su aparente desprecio por la dimensión moral del derecho. (Es decir, su dimensión moral *sustantiva*; la moral de *procedimiento* de la ley, que se resume en el concepto del imperio de la ley, está implícita en la teoría económica del derecho, como ya hemos visto.) Seguramente —se argumentará— el verdadero propósito del derecho, sobre todo el de aquellos principios fundamentales del derecho incorporados en el derecho común de Inglaterra y los Estados Unidos, es el de corregir las injusticias, reivindicando así el sentido moral. (De hecho, la sección anterior de este capítulo terminó haciendo referencia a la norma moral.)

¿Pero son realmente incongruentes la moral y la eficiencia? Para el lector atento de los capítulos anteriores será evidente el valor económico de

principios morales tales como la honestidad, la veracidad, la frugalidad, la confiabilidad (como en el cumplimiento de las promesas), la consideración por los demás, la caridad, la buena vecindad, el trabajo arduo y el evitar la negligencia y la coerción. La honestidad, la confiabilidad y el amor reducen los costos de las transacciones. La renuncia a la coerción promueve el intercambio voluntario de los bienes. La buena vecindad y otras formas del altruismo reducen los costos externos y aumentan los beneficios externos; en efecto, los economistas llaman a veces a las externalidades "efectos de la cercanía". La caridad reduce la demanda de costosos programas de beneficencia pública. El cuidado reduce el desperdicio social.[45] Es cierto que el respeto de los principios morales reduce a veces la riqueza de la sociedad: el "honor entre ladrones" ilustra esta observación. Y en el capítulo x veremos que la cartelización se ve a menudo obstruida por problemas de gorroneo que desaparecerían si los miembros de los cárteles fueran completamente desinteresados o completamente confiables. Pero, bien mirado, parecería que el respeto de los principios morales generalmente aceptados aumenta la riqueza de la sociedad en mayor medida de lo que la reduce, sobre todo si los principios están apropiadamente ordenados, de modo que el altruismo, la confiabilidad y otros rasgos instrumentales se vuelven indispensables para el bienestar social, antes que para el bienestar privado, cuando los dos discrepan.

En la medida en que el respeto de los principios morales aumente la capacidad de un individuo para maximizar sus satisfacciones, no habrá ocasión para tratar de imponer tales principios coercitivamente. Hasta cierto punto lo hace: el eslogan de "La honestidad es la mejor política" trataba de hacer un llamado al interés propio. El incumplimiento de nuestras promesas es (la mayor parte del tiempo) una mala estrategia comercial aun en ausencia de sanciones legales eficaces para el incumplimiento del contrato, porque reducirá nuestras oportunidades de transacciones en el futuro; y, en efecto, esto es más probable cuanto más débiles sean las sanciones para el incumplimiento de los contratos. La empatía con los consumidores facilitará el diseño de un producto que se venderá. Una economía de mercado se basa en el consentimiento, de modo que promueve las virtudes sociales, cooperativas. Por supuesto, también hay competencia; pero un vendedor sólo puede competir con la cooperación de sus clientes. En este sentido, la ley del mercado es diferente de la ley de la selva, a pesar del indudable paralelismo existente entre la competencia en la naturaleza y la competencia en los mercados.

El derecho común asigna costos a la violación de los principios morales que incrementan la eficiencia de una economía de mercado, pero no trata de

[45] ¿Sugiere esto alguna razón económica para que sea más probable que una persona se sienta indignada cuando la atropella un conductor descuidado que cuando es la víctima de un accidente inevitable? Su indignación podría entenderse como la consecuencia de su internalización del interés social en la eficiencia, como una virtud moral. Es poco probable que nos "indignemos" si nos alcanza un rayo.

llevar ningún principio moral hasta el límite. El derecho contractual, por ejemplo, sólo hace cumplir un limitado subconjunto de promesas; muchos incumplimientos de promesas moralmente objetables no originan ninguna causa de acción. Los costos del cumplimiento forzoso de todas las promesas excederían a los beneficios. Muchas promesas se cumplen por sí mismas (por temor a las represalias de una u otra clase) o son tan pequeñas que su cumplimiento forzoso no aumentaría el bienestar lo suficiente para justificar el costo de los remedios legales (por ejemplo, una promesa de comprarle un almuerzo a un amigo). Algunas promesas se hacen en circunstancias en que los costos del error legal exceden a los beneficios derivados del cumplimiento forzoso de la promesa en la forma en que se hizo; un ejemplo es el de un contrato que se vuelve imposible de hacer cumplir por la Ley de Fraudes porque se hizo en forma oral antes que escrita.

El compromiso del derecho común con la eficiencia se ilustra por casos tales como Reed *vs*. King[46] y Stambovsky *vs*. Ackley.[47] Éstos eran juicios de rescisión de contratos de venta de casas. En ambos casos se discutía la importancia material de la información que los vendedores no revelaron a los compradores. En Reed se trataba de que la casa había sido escenario de un homicidio múltiple diez años atrás; en Stambovsky, de que se creía ampliamente que la casa estaba habitada por fantasmas. En ambos casos, la información que retenían los vendedores era materialmente importante porque reducía el valor de las casas. El hecho de que ninguna persona "racional" pagaría menos por las propiedades a causa de estos hechos era irrelevante; los tribunales aceptaron la "subjetividad" de los valores, uno de los fundamentos de la libertad contractual y de la teoría económica moderna.

Sin embargo, los tribunales de derecho común se abstienen a veces de aplicar plenamente las implicaciones de esa teoría. No tendrían paciencia para la violación "eficiente" (discutida antes, en VII.1). No hacen cumplir los pactos suicidas, las cláusulas punitivas o los convenios racialmente restrictivos. No harán cumplir los contratos voluntarios de la esclavitud o el empleo como gladiador. La eficiencia o maximización de la riqueza es un hilo importante del tejido ético, pero no es el único.

Aun así, evaluar el costo de las políticas que pueden estar fundadas en inclinaciones morales, religiosas o emocionales fuertes siempre es una contribución útil de la economía, debido a que tales inclinaciones no son inalterables. Tengamos en cuenta la cuestión del suicidio asistido por un médico, ilegal en todos los estados, excepto en cuatro (Montana, Oregon, Vermont y Washington). Pero ¿sería tan sólida la política si resultara que el costo de la prohibición del suicidio asistido por un médico significara *más* suicidios? La economía sugiere que, de hecho, ése podría ser un costo de la prohibición. Una

[46] 193 Cal. Rptr., 130 (App., 1983).
[47] 572 N.Y. S. 2d 672 (App. Div., 1991).

persona que cree pero no tiene la certeza de que sufre una enfermedad mortal podría desear aplazar el suicidio hasta que esté segura, pero probablemente tema que el progreso de la enfermedad la vuelva incapaz de hacerlo en ese momento. En otras palabras, el costo del suicidio puede aumentar con el tiempo, lo que puede provocar un suicidio prematuro. Si se permite que un médico ayude a un paciente a suicidarse, el efecto será una reducción del costo para el paciente por suicidarse después. El resultado es hacer que la opción de esperar y ver qué pasa parezca más atractiva. De modo que más suicidas potenciales esperarán. Una fracción de éstos tendrán suerte y se recuperarán, o cambiarán de opinión por otras razones acerca de suicidarse. De ahí que, en realidad, la tasa de suicidios pueda disminuir si se permite el suicidio asistido por un médico.[48]

El principal problema ético planteado por el enfoque de la eficiencia sobre el derecho común es la discrepancia existente entre la maximización de la eficiencia y las nociones de la distribución justa de la riqueza. En la Quinta Parte examinaremos algunas teorías de la justicia distributiva. Lo importante aquí es que el derecho común es, en su mayor parte, distributivamente neutral. Las razones de esta neutralidad tienen que ver con las características institucionales de los tribunales, como veremos más adelante (véase XX.2).

Otra área de tensión involucra al paternalismo. Los economistas son menos propensos que los no economistas a apoyar las políticas paternalistas, que van desde la prohibición del suicidio asistido por un médico hasta la "guerra contra las drogas". Por otra parte, una paradoja del paternalismo es que las limitaciones cognitivas y las deficiencias de información que proporcionan el fundamento para este tipo de intervenciones pueden aquejar a los funcionarios del gobierno de igual forma o más gravemente.[49] El incentivo de un funcionario del gobierno para resistir el tirón de esas limitaciones y estar bien informado acerca de los problemas de otras personas es, en realidad, menor que el de la propia gente, porque ella paga un precio más alto por sus errores. Con todo, en los capítulos siguientes veremos que hay ejemplos de políticas paternalistas exitosas en las áreas de seguridad y de salud, como las restricciones sobre los cigarrillos y las grasas trans, y los requisitos de los cinturones de seguridad del automóvil; y existen argumentos económicos para algunas de estas políticas.

Un tema importante del paternalismo que la economía puede iluminar es el de la prohibición a los atletas de usar esteroides y otras drogas para mejorar su rendimiento. No se plantean objeciones a los atletas que mejoran su rendimiento a través de un entrenamiento mejor, más ejercicio, más práctica, o la abstención de alcohol y cigarrillos. La diferencia es que las sustancias para mejorar el rendimiento tienen efectos nocivos a largo plazo en la salud del

[48] Véase en Richard A. Posner, *Aging and Old Age*, cap. 10 (1995), una discusión más completa del suicidio asistido por un médico.

[49] Véase Edward L. Glaser, "Paternalism and Psychology", 73 *U. Chi. L. Rev.*, 133 (2006).

usuario. Esto a su vez da lugar a una externalidad en la competición atlética, ya que el uso por parte de algunos atletas desalienta el rendimiento relativo de los abstemios. Sin embargo, se plantean pocas objeciones al comportamiento autodestructivo en la búsqueda de distinción atlética, siempre y cuando el comportamiento no involucre el uso de drogas. No se critica a un defensa de fútbol americano por convertirse en un gigante de 184 kilos si lo hace sin la ayuda de drogas, a pesar de que los efectos a largo plazo del peso adicional en su salud (y en la de los jugadores con los que choca) son muy malos, y aunque su aumento de peso podría ejercer presión en otros defensas para igualar su peso, y aun cuando la mejora en la calidad de entretenimiento del fútbol americano es menor si los defensas de 184 kilos se enfrentan entre sí que entre defensas de 92 kilos. (En cambio, los asociados de un bufete de abogados que tienen demasiado trabajo aumentan la producción de la empresa.)

Aun así, la prohibición del dopaje en el deporte profesional es racional, si bien paternalista. Es un medio racional para proteger a los niños, y el "paternalismo" pierde toda connotación peyorativa cuando las personas a quienes se protege contra sí mismas son niños. Debido a que los atletas exitosos ganan sueldos elevados, porque el éxito como atleta no requiere un grado alto de inteligencia, y en virtud de que para que una trayectoria deportiva sea exitosa debe comenzar en la escuela secundaria (tal vez incluso antes, en el caso del tenis), existe una enorme competencia entre los menores para lograr el éxito deportivo. Si las sustancias que mejoran el rendimiento fueran legales, su uso por parte de los adolescentes sería generalizado, y los adolescentes carecen de la madurez suficiente para sopesar los beneficios de una trayectoria deportiva (descontados de la escasa probabilidad de que cualquier atleta adolescente dado tenga una trayectoria deportiva realmente exitosa) contra el daño a su salud a largo plazo. Desde luego, la utilización de este tipo de sustancias podría autorizarse a los atletas, pero a los menores se les prohibiría hacerlo. Sin embargo, sería difícil hacer cumplir este régimen legal, en especial dado el estatus de "modelo a seguir" de los atletas adultos a los ojos de los menores. El levantamiento de la prohibición eliminaría todo estigma del uso de estas sustancias.

Lecturas recomendadas

Becker, Gary S., "A Theory of Social Interactions", 82 *J. Pol. Econ.,* 1063 (1974).

———, "Norms and the Formation of Preferences", en Becker, *Accounting for Tastes,* 225 (1996).

Bernstein, Lisa, "Opting Out of the Legal System: Extralegal Contractual Relations in the Diamond Industry", 21 *J. Legal Stud.*, 115 (1992).

Bishop, William, "The Contract-Tort Boundary and the Economics of Insurance", 12 *J. Legal Stud.,* 241 (1983).
Coase, Ronald H., "Adam Smith's View of Man", 19 *J. Law & Econ*., 529 (1976).
Ellickson, Robert C., *Order without Law: How Neighbors Settle Disputes* (1991).
———, y Charles Dia. "Thorland, Ancient Land Law: Mesopotamia, Egypt, Israel", 71 *Chi.-Kent L. Rev.,* 321 (1995).
Glaeser, Edward L., y José Scheinkman, "Neither a Borrower nor a Lender Be: An Economic Analysis of Interest Restrictions and Usury Laws", 41 *J. Law & Econ.*, 1 (1998).
Hay, Jonathan, y Andrei Shleifer, "Private Enforcement of Public Laws: A Theory of Legal Reform", 88 *Am. Econ. Rev. Papers & Proceedings,* 398 (mayo de 1998).
Hovenkamp, Herbert, "The Economics of Legal History", 67 *Minn. L. Rev.,* 645 (1983).
Kaplow, Louis, y Steven Shavell, *Fairness versus Welfare* (2002).
Landa, Janet T., *Trust, Ethnicity, and Identity: Beyond the New Institutional Economics of Ethnic Trading Networks, Contract Law, and Gift-Exchange* (1994).
Law and Economic Development (Hans-Bernd Schäfer y Angara V. Raja [comps.], 2006).
Law, Economics, and Philosophy: A Critical Introduction with Applications to the Law of Torts (Mark Kuperberg y Charles Beitz [comps.], 1983).
Legal Institutions and Economic Development (Robert D. Cooter y Francesco Parisi [comps.], 2009).
Levmore, Saul, "Rethinking Comparative Law: Variety and Uniformity in Ancient and Modern Tort Law", 61 *Tulane L. Rev.,* 235 (1986).
Miceli, Thomas J., y Kathleen Segerson, "Punishing the Innocent Along with the Guilty: The Economics of Individual *versus* Group Punishment", 35 *J. Legal Stud.*, 81 (2007).
Parisi, Francesco, "Spontaneous Emergence of Law: Customary Law", en *Encyclopedia of Law and Economics*, vol. 5: *The Economics of Crime and Litigation*, 603 (Boudewijn Bouckaert y Gerrit De Geest [comps.], 2000).
____, "The Genesis of Liability in Ancient Law", 3 *Am. Law & Econ. Rev.,* 82 (2001).
Polinsky, A. Mitchell, y Steven Shavell, "Corruption and Optimal Law Enforcement", 81 *J. Pub. Econ.*, 1 (2001).
Posner, Eric A., *Law and Social Norms* (2000).
Posner, Richard A., *The Economics of Justice*, partes 1-2 (1981).
_____, *The Problems of Jurisprudence*, cap. 12 (1990).
Production of Legal Rules, parte 2 (Francesco Parisi [comp.], 2011).
Rubin, Paul H., *Business Firms and the Common Law: The Evolution of Efficient Rules* (1983).
Shavell, Steven, "Law *versus* Morality as Regulators of Conduct", 4 *Am. Law & Econ. Rev.*, 227 (2002).

Problemas

1. Evalúe el argumento de que los empleadores tenían un poder de negociación tan grande frente a los empleados, en el siglo XIX, que estos últimos no podían negociar las precauciones de seguridad justificadas por sus costos. Véase antes IV.8.

2. Una de las doctrinas más tristemente célebres del derecho común de los accidentes industriales era la regla del "compañero sirviente": Si un empleado era lesionado por uno de sus compañeros, no podría demandar a su empleador a menos que éste hubiera dejado de obrar con el cuidado debido en la contratación o supervisión del empleado que infligió la lesión. La regla era así una excepción a la doctrina de que el superior responde. Algunos tribunales limitaban la excepción a los trabajadores del mismo departamento de la compañía; otros la aplicaban a toda la compañía. ¿Puede imaginar una justificación económica neutral para la versión amplia o la versión estrecha de la regla del compañero sirviente? Véanse Farwell *vs.* Boston & Worcester R. R., 45 Mass. (4 Met.) 49 (1842); Pomer *vs.* Schoolman, 875 F.2d, 1262 (7th Cir., 1989).

3. ¿Cuáles serían los efectos, sobre la distribución del ingreso y la riqueza, de un fallo judicial que trasladara de los agricultores a los ferrocarriles el costo de las lesiones del ganado resultantes de no mantener al ganado fuera de las vías del ferrocarril?

4. ¿Debieran tratarse los casos de accidentes derivados de una relación contractual (por ejemplo, entre un ferrocarril y sus pasajeros, o entre un médico y sus pacientes) como un incumplimiento de contrato antes que como un cuasidelito? ¿Tiene alguna utilidad el mantenimiento de la distinción establecida entre el cuasidelito y el contrato en tales casos?

5. Analice las consideraciones morales y económicas implicadas en el hecho de permitir que los delincuentes en prisión acepten participar en experimentos médicos a cambio de una reducción de la sentencia. Véanse algunos antecedentes en National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, Staff Paper: "Prisoners as Research Subjects" (31 de octubre de 1975).

6. La Ford Motor Company decide cerrar una planta ubicada en South Bend, Indiana, y abrir una planta similar en Birmingham, Alabama. Un efecto del traslado será la disminución de los valores de la tierra en South Bend y el aumento de tales valores en Birmingham. Desde el punto de vista de la eficiencia, ¿debiera exigirse a la Ford que compensara a los terratenientes adversamente afectados?

7. Algunos estudiosos de la jurisprudencia creen que las obligaciones legales vienen primero y los derechos después; otros invierten el orden. ¿Cuál concepción es más congruente con el análisis económico?

8. En Kentucky Fried Chicken of California, Inc. *vs.* Los Angeles County Superior Court, 927 P.2d 1260 (Cal., 1997), una cliente de uno de los restaurantes de Kentucky Fried Chicken demandó a esta empresa por haberle infligido negligentemente un daño emocional. Un ladrón había amenazado con dispararle porque la cajera del restaurante se negó a abrir la caja registradora con rapidez cuando se lo exigió. La demandante sostenía que Kentucky Fried Chicken debiera haber capacitado a sus empleados para que aceptaran las exigencias de los asaltantes sin resistencia a fin de minimizar el riesgo para los clientes. Desde un punto de vista económico, ¿debiera analizarse este caso como un contrato o un cuasidelito?; y ¿quién debiera ganar?

Tercera Parte

REGULACIÓN PÚBLICA DEL MERCADO

IX. LA TEORÍA DEL MONOPOLIO

IX.1. El precio y la producción del monopolista

Observamos en el capítulo I que un vendedor no vendería a un precio menor que su costo de oportunidad, porque eso significaría sacrificar un precio mayor por uno menor. ¿Pero qué determina el límite superior de su precio? Suponiendo que su meta es la maximización de las ganancias (la diferencia entre los ingresos totales y los costos totales), su elección del precio está constreñida por la demanda de su producto y por el costo de producción.

A fin de relacionar el precio con el ingreso y, por lo tanto, con la ganancia, necesitaremos el concepto del ingreso marginal, la contribución al ingreso total de la venta de una unidad adicional. Mientras el ingreso marginal sea positivo, el ingreso total estará creciendo con la producción. Cuando el ingreso marginal baja a cero (o menos), ello significa que una venta adicional no aumentará (o disminuirá) el ingreso total.

Si se conoce la tabla de la demanda, la tabla del ingreso marginal puede derivarse matemáticamente de ella; esto se ha hecho en la gráfica IX.1.[1] El punto en que la curva del ingreso marginal se interseca con el eje horizontal marca el nivel de la producción (*q* en la gráfica IX.1) en el que se maximiza el ingreso total. El precio correspondiente a ese nivel de producción es *p*. Si el vendedor vendiera una cantidad menor, se encontraría a la izquierda de esa intersección y la producción adicional aumentaría sus ingresos totales. Si vendiera una cantidad mayor, su ingreso se encontraría en la región negativa de la curva de ingreso marginal, lo que significa que una disminución de la producción aumentaría su ingreso total.

El efecto del precio sobre la cantidad y, por ende, sobre el ingreso (precio multiplicado por la cantidad) se resume en el útil concepto de la *elasticidad*:

[1] El ingreso marginal es simplemente la tasa de cambio, o en términos matemáticos la primera derivada, del ingreso total. Y la demanda es simplemente la tabla del ingreso medio (el ingreso total dividido por la cantidad), o precio, a diversos niveles de la producción. En el caso especial donde la tabla de demanda de un producto puede aproximarse por una línea recta, como en la gráfica IX.1, la curva del ingreso marginal puede ubicarse con gran facilidad trazando una línea recta a partir de la intersección de la curva de demanda con el eje vertical hasta un punto del eje horizontal a media distancia entre el origen y la intersección de la curva de demanda con el eje horizontal. La razón de que el ingreso marginal disminuya más rápidamente que el precio es que siempre que el vendedor disminuye su precio lo disminuye en toda su producción (a menos que practique la discriminación de precios que se describirá más adelante en este capítulo); por lo tanto, el precio que obtiene para la última unidad vendida añade menos que ese precio a sus ingresos totales.

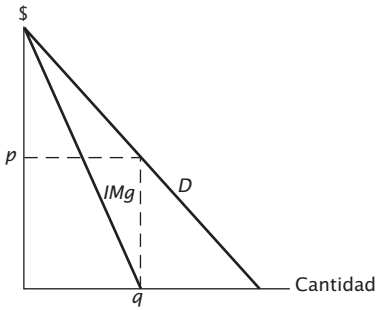
el cambio proporcional de una variable causado por un cambio proporcional en otra variable. Aquí nos interesa la elasticidad de la demanda respecto del precio o, en otras palabras, el efecto proporcional de un cambio proporcional del precio sobre la cantidad demandada. Por ejemplo, si un aumento de 1% en el precio causara una disminución de 2% en la cantidad demandada, la elasticidad de la demanda respecto del precio (o, por brevedad, simplemente la elasticidad de la demanda[2]) es –2.[3] Si la elasticidad de la demanda fuese –1, ello significaría que un aumento de 1% en el precio causaría una disminución de 1% en la cantidad demandada y, por lo tanto, que el ingreso total no cambiaría. Éste, el punto de elasticidad unitaria, es el punto en el que la curva del ingreso marginal se interseca con el eje horizontal. En todos los puntos de la curva de demanda a la izquierda de esta intersección decimos que la demanda es elástica, porque en esa región un aumento del precio conducirá a una disminución proporcionalmente mayor de la cantidad demandada y, por lo tanto, a una disminución del ingreso total. A la derecha, decimos que la demanda es inelástica porque en esta región un aumento del precio conducirá a una disminución proporcionalmente menor de la cantidad demandada y, por lo tanto, a un aumento del ingreso total.

Dado que al vendedor le interesa su ingreso neto, o ganancia, antes que su ingreso bruto, debe considerar el efecto de su elección del precio sobre su costo total tanto como sobre su ingreso total. El precio afectará a ese total determinando el número de unidades que habrán de producirse y también, si el costo marginal varía con el nivel de la producción, el costo por unidad producida. Como recordaremos, el costo marginal es el cambio de los costos totales generado por la producción de una unidad más; equivalentemente, es la adición que hace a los costos totales la última unidad producida. También hay costos fijos (el gasto en la obtención de una patente sería un buen ejemplo) —los costos que no se ven afectados por el nivel de la producción—, pero son irrelevantes para la determinación del precio y del nivel de producción porque, por definición, no son mayores, ni menores, ya sea que el monopolista cobre un precio muy elevado y produzca poco o cobre un precio muy bajo y produzca mucho.

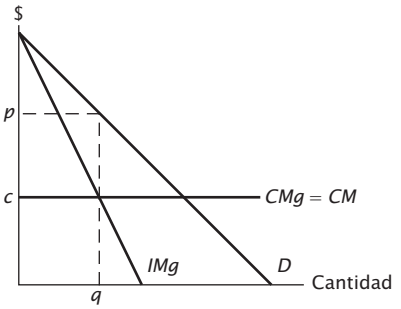
Dado que el vendedor que maximiza su ganancia expandirá su producción en la medida que una unidad adicional vendida añada más a su ingreso total que a su costo total, el nivel de producción que maximiza la ganancia es la cantidad en la que el ingreso marginal se iguala al costo marginal, q en la gráfica IX.2. A este nivel de producción, el ingreso total es igual a pq y el costo

[2] Pero la elasticidad-precio de la demanda no es el único tipo de elasticidad de la demanda. Véase antes I.1, notas 3, 4 y el texto que la acompaña (una discusión implícita de la elasticidad-ingreso de la demanda).

[3] La razón del signo menos es que un cambio en el precio tiene el efecto contrario sobre la cantidad; por ejemplo, un aumento en el precio provoca una disminución en la cantidad.



GRÁFICA IX.1

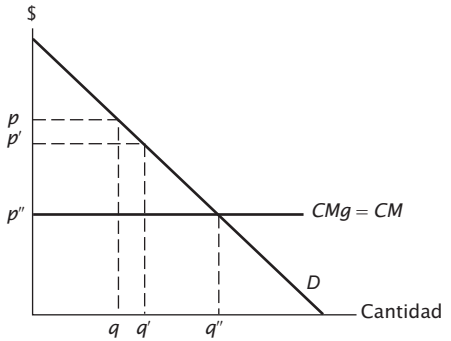


GRÁFICA IX.2

total (costo medio multiplicado por la cantidad) es igual a cq.[4] Adviértase que si el nivel de producción fuese menor, la ganancia sería menor también, porque el vendedor se encontraría a la izquierda de la intersección, donde la producción adicional añadiría más al ingreso total que al costo total. Con una producción mayor, la ganancia sería menor también. El vendedor se encontraría a la derecha de la intersección, en la región donde cada unidad vendida añade más al costo total que al ingreso total (es decir, donde el costo marginal supera al ingreso marginal).

Otro nombre para p en la gráfica IX.2 es el precio de monopolio, porque es el precio que cobraría una empresa que no tenga competencia, o que sienta temor de la competencia. La competencia haría insostenible a ese precio. Supongamos que A, inicialmente el único vendedor del producto (adminículos, por supuesto) cuyas curvas de costo y de demanda se muestran en las gráficas IX.1 y IX.2, establece un precio de p y vende q adminículos, y que otros vendedores pueden producir y vender adminículos al mismo costo que A. Uno de ellos, B, atraído por las grandes ganancias de A ($pq - cq$), decide fabricar y vender adminículos, también al precio p. La gráfica IX.3 muestra el resultado. Si B vende la cuarta parte de los adminículos que vende A (la producción de B es $q' - q$), la cantidad total de adminículos en el mercado (q') es una cuarta parte mayor que antes. Los consumidores no pagarán p por la mayor cantidad de adminículos, sino sólo p'. Es posible que A y B disminuyan su nivel de producción; pero C, D, E y otros están esperando para entrar al mercado a fin de capturar una parte de las ganancias que atrajeron a B. Las entradas continuarán hasta que el precio baje a p'', donde es justamente igual al costo de oportunidad de la producción de adminículos adicionales. Hasta que se llegue a ese punto, los recursos destinados a la fabricación de adminículos ganarán más que su costo de oportunidad y, por lo tanto, más de lo que podrían ganar en mercados competitivos. La producción de admi-

[4] Para mantener simples las cosas, se supone que los costos marginales son constantes y que la empresa no tiene costos fijos, es decir, sólo tiene costos variables. Bajo estos supuestos, el costo marginal es igual al costo medio.



GRÁFICA IX.3

níclos no se expandirá más allá de q'' (correspondiente a p'') porque entonces tales recursos estarían ganando menos que su costo de oportunidad.[5]

Podría parecer que la posibilidad de la entrada hace del monopolio un concepto académico. Dado que el beneficio de un monopolio es mayor que uno competitivo, los vendedores en mercados competitivos se moverán hacia un mercado en el que se esté cobrando un precio de monopolio. Para obtener algunos de los beneficios del monopolio en ese mercado, el participante nuevo debe vender su producto, y sus ventas incrementarán la producción del mercado, provocando una disminución en el precio. De este modo, la mayoría de los monopolios contienen la semilla de su propia destrucción. Pero la entrada requiere a veces mucho tiempo, o está prohibida, o el de nuevo ingreso no puede producir a un costo tan bajo como el de la empresa existente, al menos al principio (debido a la curva de aprendizaje; véase XI.2 después). Pero la rapidez con la que las nuevas empresas entran a un mercado en el que se está cobrando el precio de monopolio es crítica. Si el monopolista tiene una patente u otro monopolio legal, la rapidez podría ser cero y el precio de monopolio persistirá hasta que termine el monopolio legal.

Algunas veces el monopolio persistirá sin ninguna barrera legal a la entrada. Los costos del monopolista pueden ser a tal punto más bajos que aquellos de cualquier participante nuevo que el precio de monopolio es menor al precio que tendría que cobrar un participante nuevo para poder cubrir sus costos. (Ésa puede ser la razón de que exista el monopolio.) O puede ser que el precio de monopolio, aunque sea más alto que los costos del de nuevo ingreso, no tenga ningún atractivo porque el participante potencial sabe que si entra al mercado el monopolista puede cobrar con facilidad un precio remunerativo menor a los costos del participante nuevo, siendo el monopolista el productor más eficiente. El monopolio también puede ser una

[5] Bajo la competencia perfecta, la elasticidad de la demanda de cada empresa es infinita: la elevación del precio más mínima hará que la cantidad demandada disminuya a cero. Por lo tanto, la curva de la demanda es horizontal.

condición duradera del mercado debido a que un solo vendedor puede proveer a todo el mercado a un costo menor que múltiples vendedores; a esto se le llama "monopolio natural" (véase XIII.1 más adelante).

Pero incluso si los costos del participante nuevo son los mismos que los del monopolista, de esto no se sigue que la amenaza de la entrada siempre disuada de hacer el cobro del precio de monopolio. Dado que el costo está negativamente relacionado con el tiempo[6] (sería más costoso construir una planta de acero en tres meses que en tres años), la entrada *inmediata* a un mercado monopolizado a costos comparables con los del monopolista a menudo será imposible. Por lo que habrá un periodo durante el cual puedan obtenerse los beneficios del monopolio, aun cuando no existan barreras a la entrada en el sentido de una desventaja de costo para un participante nuevo.[7] Y la curva de aprendizaje que mencionamos antes podría dar al monopolista una ventaja de costo temporal sobre un participante nuevo justo por el solo hecho de haber estado más tiempo en el mercado.

El monopolio puede anticiparse y la entrada se puede planear para que coincida con su creación… pero tal vez no; el monopolista aspirante (o cártel) intentará ocultar sus intenciones. Además, los planes para entrar del futuro participante, si se entera de ello el monopolista aspirante, podrían disuadirlo de formar el monopolio. En ese caso, los gastos del futuro participante en la preparación para entrar serán en vano porque se supone que sólo fue la expectativa de que haría un monopolio lo que motivó la entrada planeada.

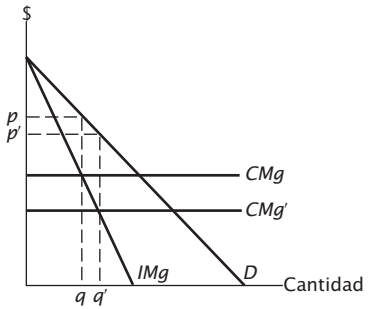
IX.2. Efecto de los cambios del costo o de la demanda sobre el precio de monopolio

Veamos lo que ocurre con el precio y la producción óptimos del monopolio si los costos del monopolista aumentan o disminuyen, o si cambia la demanda. Si los costos bajan (a menos que se trate de costos fijos), el precio óptimo del monopolio bajará y la producción aumentará. Recuérdese que el monopolista deja de expandir su producción en el punto donde se intersecan sus curvas de ingreso marginal y de costo marginal. Si la curva de costo marginal baja, la curva del ingreso marginal se encontrará ahora por encima de ella, y el monopolista expandirá su producción hasta que esa curva se interseque de nuevo con la curva del costo marginal. Esto se muestra en la gráfica IX.4.

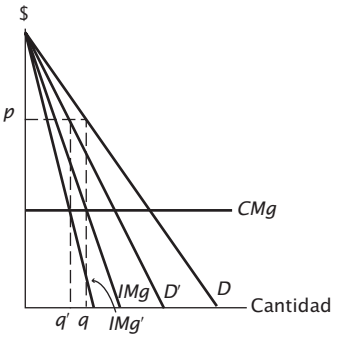
Supongamos que la demanda disminuye. Es decir, a cualquier precio dado los consumidores comprarán menos del bien en un porcentaje igual (en la gráfica IX.5, cerca de 25% menos). Si los costos marginales son constantes, como en ese diagrama, el monopolista no cambiará su precio. El costo

[6] Armen A. Alchian, "Costs and Outputs", en *Readings in Microeconomics*, 159, 165 (William Breit y Harold M. Hochman [comps.], 2ª. ed., 1971).

[7] Véase sobre las barreras a la entrada después en X.10.



GRÁFICA IX.4



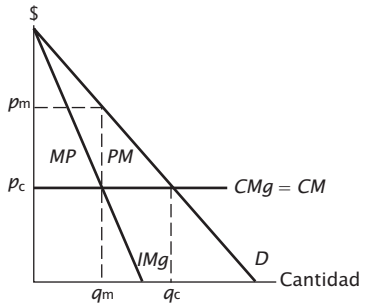
GRÁFICA IX.5

marginal no cambia. Tampoco cambia la sensibilidad (elasticidad) de la cantidad demandada a un cambio en el precio: aunque los consumidores comprarán menos a cualquier precio, el cambio proporcional de la cantidad demandada, generado por un cambio del precio, permanecerá constante. Por lo tanto, el monopolista cobrará el mismo precio que antes, pero venderá menos porque la demanda es menor. Esto ilustra el punto importante, del que volveremos a ocuparnos en el capítulo siguiente, de que el precio de monopolio depende sólo de la elasticidad de la demanda y de los costos marginales.

Si la demanda baja y los costos marginales del monopolista no son constantes, el precio del monopolio cambiará. Una curva de costo marginal no horizontal implica que el costo marginal es distinto a niveles diferentes de la cantidad producida. Dado que un cambio de la demanda alterará el nivel óptimo de la producción del monopolista, su costo marginal cambiará y, por lo tanto (según la discusión anterior), también cambiará su precio.

IX.3. CONSECUENCIAS DE EFICIENCIA DEL MONOPOLIO; LA PÉRDIDA DE PESO MUERTO Y LA BÚSQUEDA DE RENTA

Una revisión más de la gráfica IX.3 mostrará que la producción es menor bajo el monopolio que bajo la competencia: q en lugar de q''. Esto es así porque el precio de monopolio hace que algunos consumidores compren otros productos que se vuelven más atractivos por el aumento del precio del monopolio. La sustitución implica una pérdida de valor. Esto puede verse muy fácilmente suponiendo que para cada uso del producto monopolizado hay un producto sustituto idéntico pero cuya producción es simplemente más costosa, y, por lo tanto, aunque se vende al precio competitivo, su precio es mayor que el precio que se fijaría al producto monopolizado si se vendiera a su precio competitivo, pero menor que el precio de monopolio. El efecto del monopo-



GRÁFICA IX.6

lio consiste en hacer que algunos consumidores satisfagan sus demandas cambiándose a bienes cuya producción es más costosa para la sociedad que la del bien monopolizado. El triángulo marcado *PM* (por la pérdida del peso muerto) se aproxima a ese costo en la gráfica IX.6. Imaginemos que el precio aumenta gradualmente de p_c a p_m; en cada aumento, los consumidores se desplazan a sustitutos más costosos, y el último consumidor desplazado compra un producto cuya producción cuesta infinitesimalmente menos que el precio de monopolio.[8]

El concepto de la pérdida de peso muerto ilustra muy bien la distinción destacada en el capítulo anterior entre las concepciones económicas y utilitarias del bienestar. La conclusión de que *PM* en la gráfica IX.6 es un costo social neto se basa en el supuesto de que un dólar vale lo mismo para consumidores y productores o, dicho de otro modo, que los cambios en la distribución del ingreso y de la riqueza no se considerarán al evaluar los costos sociales. *MP,* la transferencia de riqueza de los consumidores a los productores, gene-

[8] Sin embargo, adviértase la dependencia de este análisis respecto del supuesto de que el sustituto se vende a un precio competitivo. En caso contrario, el monopolio podría prevenir, en lugar de crear, una pérdida de peso muerto. Supongamos que los botones de cuero están monopolizados y se venden a un precio de 10 centavos, aunque el costo de su producción es de sólo 6 centavos. Los botones de plástico, el sustituto más cercano de los botones de cuero, se venden a 8 centavos cada uno, por lo que algunas personas sustituyen los de cuero por aquéllos. Si el precio de un botón de plástico es igual a su costo, esta sustitución es ineficiente. Se está comprando un producto cuya producción cuesta 8 centavos en lugar de uno cuya producción cuesta sólo 6 centavos.

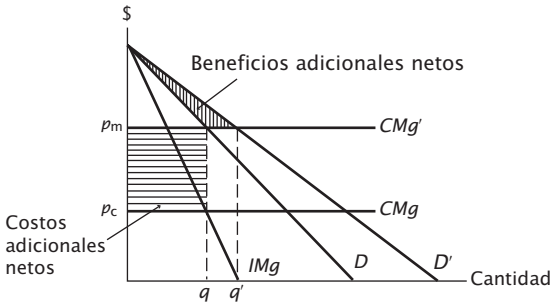
Pero supongamos ahora que los botones de plástico no se venden a un precio igual a su costo: su precio de 8 centavos es un precio de monopolio; su producción sólo cuesta 2 centavos. Cuando los botones de cuero se venden al precio de monopolio de 10 centavos, el consumidor comprará más botones de plástico a pesar de *su* precio de monopolio, y esto es eficiente porque la producción de botones de plástico es menos costosa. Si ahora se termina el monopolio de los botones de cuero, el precio bajará hasta el nivel competitivo de 6 centavos, y los consumidores empezarán a sustituir los botones de plástico por los de cuero. Ésta es una sustitución ineficiente, suponiendo que los botones de cuero no valen 4 centavos (6 – 2 centavos) más que los botones de plástico. Sin embargo, resulta dudosa la importancia empírica de este tipo de problema (el problema del "segundo mejor"). Richard Schmalensee, *The Control of Natural Monopolies,* 30-31 (1979).

rada por el aumento del precio del nivel competitivo al de monopolio, se trata como inexistente; la pérdida de los consumidores es la ganancia de los productores. El costo social del monopolio sería mayor que *PM* si se pensara que una pérdida de un dólar para los consumidores rinde menos de un dólar de beneficio para los productores, y sería menor que *PM* si se supusiera lo contrario. Si, en promedio, los consumidores son más pobres que los productores (más precisamente, que los propietarios de los productores) y, por lo tanto, debido a una declinación en la utilidad marginal del ingreso, un dólar vale más para el consumidor promedio que para el productor promedio, el monopolio reducirá la utilidad aun cuando no haya efecto alguno en la producción y, por tanto, no haya una pérdida en el peso muerto. Este punto se ignora en muchas discusiones del costo social del monopolio.

La transferencia de riqueza de los consumidores a los productores, generada por el precio de monopolio, es un resultado de la conversión del excedente del consumidor en excedente del productor. El excedente del consumidor es el área bajo la curva de demanda por encima del precio competitivo; en la gráfica IX.6 es el triángulo cuya base es la línea designada *CMg* = *CM*. *MP* es la parte del triángulo que se transforma en excedente del productor. *PM* es la parte que se pierde. El excedente del consumidor es, pues, una medida del valor agregado que los consumidores asignan a un producto por encima del precio que pagan por él. A medida que el precio aumenta, los consumidores que no valoren mucho el producto se desplazan hacia sustitutos, mientras que quienes se quedan con el producto obtienen menos valor neto de sus compras porque están pagando un precio mayor.

Aunque *MP*, la parte del excedente del consumidor que se convierte en excedente del productor, parece un pago de transferencia puro, puede ser una fuente de costo social aun cuando no se asignen ponderaciones distributivas a las transferencias de consumidores a productores. Supongamos en la gráfica IX.7 que el mercado contiene muchas empresas pero el gobierno fija el precio en p_m (precio de monopolio) y prohíbe nuevas entradas al mercado.[9] Dado que cada venta adicional que pueda hacer un vendedor le generará una ganancia monopólica de $p_m - p_c$, cada vendedor tendrá un incentivo para aumentar su participación en el mercado. No lo puede hacer disminuyendo el precio porque eso está prohibido, de modo que tratará de volver más atractivo su producto que el de sus competidores en otras formas: aumentando su calidad, proveyendo mejor servicio, etc. El proceso de competencia en variables distintas del precio continuará, en ausencia de algún acuerdo que limite tal competencia, hasta que los costos marginales de los vendedores aumenten al nivel del precio fijo, en cuyo punto los gastos adicionales en el mejoramien-

[9] Si aun cuando el precio está fijo no se impide la entrada de nuevos participantes, las empresas ofrecerán una mayor calidad o un mejor servicio para atraer las ventas de las empresas existentes, y el proceso continuará hasta que los costos marginales del último participante nuevo igualen el precio (fijo) del mercado.



GRÁFICA IX.7

to del producto generarían una pérdida. Aunque este proceso competitivo aumentará el valor del producto para el consumidor (es decir, la demanda aumentará), los costos de las mejoras podrían superar a los beneficios del consumidor, generando una pérdida social neta como en la gráfica IX.7. Por supuesto, esto es sólo nuestra vieja amiga: la búsqueda de renta (véase antes, por ejemplo, III.2).

Hasta la desregulación de la industria aeronáutica a principios del decenio de 1980, la Junta de Aeronáutica Civil (CAB) impedía la competencia de precios entre las aerolíneas que regulaba y la creación de nuevas aerolíneas troncales. Como resultado, los precios cobrados por las aerolíneas excedían los niveles competitivos. Aun así, la industria no tenía ganancias monopólicas y en efecto estaba periódicamente al borde de la quiebra. Las ganancias monopólicas potenciales se habían transformado en costos por efecto de una competencia vigorosa en variables distintas del precio, especialmente la programación de vuelos excesivamente frecuentes en relación con la demanda, lo que se traducía en baja utilización de la capacidad y, por lo tanto, en altos costos por pasajero transportado.[10]

Cuando el gobierno no trata de impedir la competencia de precios y la entrada nueva, el problema de la competencia dispendiosa por el servicio es menos agudo. Pero si una empresa o un cártel tiene, o cree que puede obtener, beneficios monopólicos, el mercado está en desequilibrio e inducirá a las empresas a gastar recursos en la formación y el mantenimiento de monopolios y cárteles y, una vez formados (en el caso de un cártel), en el acaparamiento de la mayor porción posible de las ventas del mercado mediante la competencia en variables distintas del precio. Estos recursos se desperdiciarán (en gran medida) desde un punto de vista social.

[10] Véanse, por ejemplo, George W. Douglas y James C. Miller, III, "The CAB's Domestic Passenger Fare Investigation", 5 *Bell J. Econ. & Mgmt. Sci*, 205 (1974); *Airline Deregulation: The Early Experience* (John R. Meyer y Clinton V. Oster Jr. [comps.], 1981); Thomas Gale Moore, "US Airline Deregulation: Its Effects on Passengers, Capital, and Labor", 29 *J. Law & Econ.*, 1 (1986).

Adviértase el estrecho paralelo con el análisis de los costos de las transferencias coercitivas puras, como el robo, en el capítulo VI y especialmente el capítulo VII. El robo es también "sólo" un pago de transferencia; la pérdida de la víctima es la ganancia del ladrón. La objeción económica contra el robo no implica a la transferencia como tal sino el hecho de que la oportunidad de obtener tales transferencias inducirá un gasto de recursos en el robo y en la prevención del robo. Hay una objeción económica idéntica contra el monopolio, que es distinta del efecto del precio de monopolio sobre la producción: un efecto que no tiene contrapartida en el robo.

Pero no debemos suponer que *todas* las ganancias monopólicas potenciales, u otras rentas, se conviertan necesariamente en costos. Si la expansión de una actividad se topa con rendimientos fuertemente decrecientes (en otras palabras, si el costo marginal está aumentando considerablemente), el actor dejará de expandirse mucho tiempo antes de que haya gastado recursos iguales a las ganancias potenciales máximas derivadas de esa actividad. Una empresa racional decidida a obtener un monopolio, en lugar de presupuestar para este proyecto una cantidad igual a todo el valor presente descontado de sus ganancias monopólicas, dejará de gastar en el punto en el que un dólar adicional (en publicidad, competencia de servicios, abogados, etc.) genere menos de un dólar en ganancias esperadas.[11] Por otra parte, podría haber una duplicación de gastos de varios monopolistas potenciales, como en nuestro ejemplo de la búsqueda de tesoros en III.2: esta duplicación debe reconocerse también en el cálculo del costo del monopolio.

IX.4. Otras objeciones económicas contra el monopolio; origen de la innovación y la minimización del costo

Un monopolista podría ser capaz de desplazar la curva de demanda de su producto hacia fuera —lo que conduciría a la producción excesiva en vez de una producción insuficiente— disfrazando el precio o la calidad de su producto (véase antes IV.7) o logrando que el gobierno disminuya la oferta de productos rivales o aumente la demanda de productos complementarios.[12]

[11] Por ejemplo, supongamos que una unidad de monopolización cuesta 1 dólar y crea una ganancia esperada del monopolio de 10 dólares, una segunda unidad costaría 2 dólares y aumentaría la ganancia esperada del monopolio a 15 dólares, una tercera unidad costaría 8 dólares y aumentaría la ganancia esperada del monopolio a 25 dólares, pero una cuarta unidad costaría 16 dólares y aumentaría la ganancia esperada del monopolio a 30 dólares. Entonces la empresa se detendría en tres unidades porque la cuarta unidad compraría sólo 5 dólares en ganancias adicionales esperadas del monopolio pero costaría 16 dólares. El costo marginal de la empresa a ese nivel de producción es 8 dólares y su costo total es de sólo 11 dólares, que es considerablemente menor que su ganancia esperada del monopolio de 25 dólares. En este ejemplo sólo se ha transformado en costo cerca de 40% de la ganancia esperada.

[12] Dos productos son complementarios cuando una disminución del precio de uno de ellos conduce a un aumento de la cantidad demandada del otro. Más adelante discutiremos la capaci-

De mayor importancia es que la preocupación de que el monopolio disminuya el incentivo de una empresa para innovar y para minimizar sus costos. Esto plantea una cuestión importante, dado el potencial de la innovación para destruir los beneficios de la competencia estática. Es famoso el argumento de Joseph Schumpeter en el sentido de que el bienestar económico se maximiza a lo largo del tiempo a resultas de una sucesión de monopolios, un proceso de competencia dinámica que él llamó "el huracán de la destrucción creativa".[13] Cada monopolista le arrebata el control del mercado a su predecesor mediante innovaciones que reducen los costos o mejoran los productos y le otorgan a su vez un monopolio temporal; éste le permite recuperar el gasto de su innovación con una ganancia suficiente para compensarlo por el riesgo de fracaso, que es considerable en el caso de la innovación debido a su incertidumbre inherente. De hecho, como Schumpeter señaló, la sucesión de monopolios podría aumentar más el bienestar social de lo que lo haría la competencia estática. Compárese un mercado donde el costo marginal es 10 dólares y el precio es también 10 dólares —en otras palabras, un mercado en equilibrio competitivo— con un mercado por lo demás idéntico donde el costo marginal es 6 dólares y el precio es 8 dólares porque entró al mercado una empresa con un procedimiento nuevo para reducir el costo de producción que desplazó a las empresas existentes mediante la fijación de precios por debajo de sus costos (pero por encima de sus costos más bajos) y, habiendo logrado así la constitución de un monopolio, está obteniendo un rendimiento monopólico y continuará haciéndolo hasta que la desplace una empresa todavía más innovadora. Los consumidores, y la sociedad en conjunto, se encuentran en mejor situación bajo el segundo escenario, donde la perspectiva de ganancias monopólicas crea el incentivo indispensable para buscar la innovación riesgosa, socialmente benéfica.

Pero la teoría de Schumpeter no dice nada acerca de la estructura de mercado óptima para *iniciar* el huracán de la destrucción creativa. No está claro si es mejor tener cierto número de competidores, cada uno de los cuales trata de convertirse en un monopolista mediante la innovación, o tener en cambio un monopolista único que trata de aumentar sus ganancias monopólicas, o bien (lo que podría ser lo mismo) impedir la competencia futura mediante la innovación. Aunque ni la teoría ni los estudios empíricos contestan este interrogante de manera definitiva,[14] el peso del argumento está contra el mono-

dad de las empresas para obtener asistencia gubernamental en la supresión de la competencia. Véase especialmente el capítulo XX. Una industria competitiva, actuando a través de una asociación industrial, podría ser capaz de obtener la misma clase de asistencia gubernamental que un monopolista. Éste es un ejemplo de la proposición más general de que las empresas independientes pueden obtener a menudo, mediante la coordinación, los mismos resultados que el monopolio. Véase después X.1.

[13] Véase Joseph A. Schumpeter, *Capitalism, Socialism, and Democracy,* 81-106 (1942).

[14] Véanse Wesley M. Cohen y Richard C. Levin, "Empirical Studies of Innovation and Market Structure", en *Handbook of Industrial Organization*, vol. 2, p. 1060 (Richard Schmalensee y Ro-

polio. Todo lo que puede decirse a su favor es que el monopolista puede ser capaz de apropiarse más de los beneficios de una innovación, mientras que, en un mercado competitivo, el innovador beneficiará a sus competidores así como a sí mismo. Sin embargo, esto pasa por alto la protección de patentes y otras ventajas del que "mueve primero" (la primera empresa que desarrolla una innovación puede producirla y comercializarla de manera más eficiente que los que copian y crear lealtad hacia la marca que sería difícil de desplazar para los recién llegados) que podrían brindar al innovador la protección suficiente en contra de la apropiación de su innovación por los competidores para que pueda recuperar el costo de la innovación con un beneficio sustancioso. En contra del monopolio como el mejor entorno para la innovación, puede argumentarse que:

1. El monopolista tiene menos que ganar de la innovación de lo que ganaría una empresa competitiva. Ya se ha apropiado de gran parte del excedente del consumidor disponible; la empresa competitiva que podría convertirse en un monopolista mediante la innovación no lo ha hecho. Este punto se aplica tanto a la innovación de productos como a la innovación de procesos. La innovación de productos permitirá que el monopolista se apropie de un poco más del excedente del consumidor. La innovación de procesos (una forma más barata de hacer el mismo producto) reducirá un poco los costos del monopolista. Cualquier tipo de innovación podría posibilitar que una empresa tome el control de todo el mercado y, de ese modo, se apropie de una gran muestra de excedente del consumidor al elevar su precio del nivel de competencia al nivel del monopolista.

2. El monopolista también tiene menos que perder que la empresa competitiva de la *falta* de innovación. La empresa competitiva puede ir a la quiebra, y la quiebra es más costosa para los directores de una empresa que el hecho de simplemente dejar de obtener grandes ganancias, debido a la aversión al riesgo de los administradores corporativos (véase el capítulo XV). Más de la riqueza de los directores estará atada en una empresa única, y si la empresa se declara en quiebra también perderán la capacidad de ganancia específica de esa empresa (el "capital humano específico de la empresa"). El economista inglés John Hicks dijo la famosa frase "el mejor de todos los beneficios del monopolio es una vida tranquila".[15] Al decir eso negaba la equivalencia de los costos extra o gastos de desembolso y los costos de oportunidad. El monopolista que no logra mantener su barco bien amarrado está regalando dinero; y sabemos que (éste es otro dicho favorito de los economistas) nadie deja de recoger un billete de un dólar que ve tirado en la acera: eso sería irracional porque el costo de tiempo de recoger el dólar debe ser menor a un

bert D. Willig [comps.], 1989); F. M. Scherer, *Industrial Market Structure and Economic Performance*, caps. 15-16 (3ª ed., 1990).

[15] S. R. Hicks, "Annual Survey of Economic Theory: The Theory of Monopoly", 3 *Econometrica* 1, 8 (1935).

dólar. Pero el realismo acerca de la gestión de las empresas, combinado con la incertidumbre, la aversión al riesgo (una correlación de la utilidad marginal decreciente del ingreso) y la aversión a la pérdida (todas analizadas antes en I.1), sugieren que Hicks puede haber tenido razón y, por consiguiente, que el costo de oportunidad no es equivalente al costo extra. Si la gestión está dirigiendo un negocio rentable, o una división rentable de una empresa, existe una reticencia natural al cambio, porque existe el riesgo de fallar y se culpará a la gestión. Y si se está haciendo todo bien, es poco probable que la posibilidad de hacerlo aún mejor prevalezca por encima de la oportunidad de hacerlo peor. (Éste es un ejemplo de aversión a la pérdida, discutido antes en I.4.)

Además, el cambio produce malestar en el lugar de trabajo. Los trabajadores temen por sus empleos y podrían verse presionados para trabajar más, y podrían reducirse los beneficios o adoptarse otras medidas para economizar en los costos laborales, creando tensiones y hostilidades que son costos reales, aunque no monetarios, para la gestión. En una organización, como resultado de la renuencia al cambio, la desaceleración se acumula en épocas de prosperidad. Cuando la organización atraviesa tiempos difíciles, debido a una mayor competencia o a un cambio en las preferencias del consumidor, la presión para reducir los costos crece, y la desaceleración acumulada en épocas de prosperidad brinda una oportunidad para economizar que ya no es posible ignorar. De este modo, los costos extra se reducen en la adversidad, pero se permite la acumulación de los costos de oportunidad en el marco de la prosperidad.

3. Las empresas difieren en su capacidad para innovar. Es más probable que un mercado con varias empresas contenga por lo menos un innovador por encima del promedio que un mercado que contenga una sola empresa y, si es así, el mercado competitivo será más innovador que el mercado monopolista.

4. Cuando existe incertidumbre sobre la mejor manera de lograr un objetivo, podría ser más eficiente buscar diferentes caminos que tratar de averiguar desde el principio del proceso innovador qué camino es el más prometedor. La analogía es con la selección natural, en la que la lucha por la supervivencia (análoga a la competencia económica) selecciona a aquellas formas mutantes mejor adaptadas a un entorno cambiante e incierto. Si, de modo similar, la innovación es un proceso en el que el mercado hace una selección de entre los distintos enfoques cuya promesa relativa no se puede evaluar con anticipación, entonces es probable que mientras haya más empresas rivales que sigan caminos diferentes para lograr algún descubrimiento ingenioso, éste se logre más rápido.[16] Pero ¿no podría una empresa monopolista simular el proceso al hacer que diferentes divisiones de la empresa tomen caminos diferentes? En principio, sí, pero es probable que sea diferente en la práctica. Es probable que los distintos caminos de la invención requieran culturas organi-

[16] Richard R. Nelson y Sidney G. Winter, *An Evolutionary Theory of Economic Change* (1982).

zacionales diferentes, y para una empresa es difícil internalizar diferentes culturas porque eso dificulta controlarla.

5. Es difícil evaluar la gestión de un monopolista, porque no existe referencia alguna: una empresa similar que pudiera tener costos más bajos o participar en una innovación mayor porque tuviera mejores administradores, o al menos administradores que fueran agentes más fieles a los accionistas.[17] Esto sugiere, a su vez —y existe cierta evidencia—, que un mercado competitivo tenderá a ser más exitoso que un mercado monopolizado en lo tocante a la minimización de los costos existentes (por oposición a la innovación de procesos o productos).[18] Como se señaló antes, la competencia provee un punto de referencia para la evaluación del desempeño de un administrador.

El monopolio podría alterar las preferencias de los consumidores por la calidad o la variedad.[19] El monopolista podría utilizar la calidad o la variedad para tratar de identificar la elasticidad de la demanda de los distintos grupos de clientes con el fin de practicar la discriminación de precios de forma más efectiva. Por ejemplo, en la industria de la transmisión un monopolista podría atender más a los gustos de una minoría (es decir, ofrecer una mayor variedad de programación) que las empresas competitivas porque todas las empresas estarían apuntando hacia el espectador o el oyente promedio (un efecto de congestión) y haciendo caso omiso de los costos para cada una; el monopolista internaliza esta externalidad.

Dado que los precios de monopolio pueden resultar de la colusión entre las empresas rivales y no de la ausencia del mercado de cualquier otra empresa, debemos considerar los incentivos para innovar y reducir costos en un mercado que contiene más de una empresa pero que está efectivamente cartelizado y en el que la entrada de nuevas empresas se encuentra obstruida. Al revés de lo que ocurre con un monopolista de una sola empresa, el miembro de un cártel debe considerar la posibilidad de que el cártel se derrumbe, lanzándolo de regreso a la arena competitiva donde se encontrará en peligro mortal si sus costos superan a los de las demás empresas de su mercado. Aun si el cártel dura mucho tiempo, los miembros que tengan los costos más bajos presionarán en favor de una reducción del precio del cártel, ya que entre menores sean los costos marginales de una empresa, en igualdad de todas las demás circunstancias, será menor el precio que maximice sus ganancias (véase antes IX.2). Su presión tendrá cierto efecto, ya que pueden amenazar creíblemente con superar a los otros miembros en la competencia en caso de que el cártel se derrumbe.

[17] Véase Joseph Farrell, "Monopoly Slack and Competitive Rigor: A Simple Model", en *Readings in Games and Information*, 169 (Eric Rasmusen [comp.], 2001).

[18] Véanse Stephen J. Nickel, "Competition and Corporate Performance", 104 *Journal Pol. Econ.,* 724 (1996); Michael Gort y Nakil Sung, "Competition and Productivity Growth: The Case of the U.S. Telephone Industry", 37 *Economic Inquiry,* 678 (1999).

[19] Véase, por ejemplo, Michael Musa y Sherwin Rosen, "Monopoly and Product Quality", 18 *J. Econ. Theory,* 301 (1978).

Una consideración adicional: los miembros de un cártel raras veces encuentran viable ponerse de acuerdo para no competir en *todas* las dimensiones de la posible rivalidad. De ordinario sólo se fijan el precio y otras condiciones de la venta, y las empresas quedan en libertad para competir en otros sentidos. Como hemos visto, un efecto de la colusión para fijar los precios cuando no está restringida la competencia en variables distintas del precio es la canalización de las ganancias generadas por tal fijación de los precios hacia los gastos en la competencia en variables distintas del precio, incluyendo los mejoramientos del producto y el servicio. Si acaso, por lo tanto, esperaríamos que la cartelización aumentara el incentivo para inventar por comparación con lo que ocurriría en un mercado de precios competitivos… y aun para impulsar la invención más allá del punto óptimo. Sin embargo, supongamos que la entrada de empresas nuevas es factible. Esto reduciría el incentivo para inventar si la empresa nueva pudiera aprovecharse gratuitamente de la actividad inventiva de la competencia, ya que la empresa nueva no habría contribuido a los costos fijos de la invención.

IX.5. Discriminación de precios

Hasta ahora hemos supuesto que el monopolista vende su producto a un solo precio. Este supuesto se justifica de ordinario. Si un monopolista vendiera a dos precios diferentes, los compradores a quienes se cobrara el precio menor revenderían a los compradores a quienes se cobre el precio mayor, obteniendo una ganancia fácil. Tal reventa (conocida como arbitraje) haría imposible que el vendedor mantuviera precios diferentes. Pero el arbitraje no es siempre viable. Es posible que el producto sea imposible de almacenar para su reventa futura (como ocurre en el caso de muchos servicios), o que existan restricciones contractuales para la reventa. Un monopolista que no se enfrente al daño del arbitraje tenderá probablemente a fijar precios diferentes para compradores diferentes, dependiendo no de los costos de la venta a ellos —suponiendo que los costos son los mismos, como bien pueden serlo—, sino de la elasticidad de sus demandas del producto del monopolista. Ésta es la discriminación de precios.

La pendiente negativa de la curva de demanda implica que algunos consumidores estarían dispuestos a pagar precios muy por encima del costo por al menos algunas unidades de la producción del monopolista, mientras que otros estarían dispuestos a pagar precios sólo ligeramente (o no) mayores que ese costo. De modo que no hay un solo precio que extraiga el valor total que algunos consumidores obtienen del producto sin sacrificar las ganancias menores pero todavía positivas al que pudieran venderse unidades adicionales. Idealmente, al monopolista le gustaría negociar por separado (pero sin costo) con cada consumidor por cada unidad. Entonces no rechazaría nunca a un cliente

dispuesto a pagar un precio igual al costo, y su producción sería idéntica a aquella situada bajo la competencia, pero sus ganancias serían mucho mayores. Pero los costos de transacción de la discriminación perfecta de los precios (llamada también "de primer grado") son prohibitivos. De ordinario, lo mejor que puede hacer el monopolista discriminador consiste en dividir a sus clientes en unos cuantos grupos y fijar un solo precio (si bien diferente) para cada grupo.

Es indeterminado el efecto sobre la producción de la discriminación de precios imperfecta (a la que los economistas llaman de tercer grado).[20] Supongamos que el precio monopólico único maximizador de la ganancia es 10 dólares; pero, en lugar de cobrar ese precio, el monopolista clasifica a los compradores en dos grupos y cobra 5 dólares a los miembros de uno de ellos y 20 dólares a los miembros del otro. El monopolista gana ventas —a los miembros del primer grupo que no pagarán 10 dólares pero están dispuestos a pagar 5—. Pero también pierde ventas —a los miembros del segundo grupo que están dispuestos a pagar 10 dólares pero se niegan a pagar 20—. Las ventas perdidas podrían superar a las ventas ganadas. La producción del monopolista podría ser menor aunque sus ganancias fuesen mayores.

Ni siquiera la discriminación perfecta de los precios eliminaría la objeción económica contra el monopolio porque, ahora que son mayores las ganancias potenciales del monopolio,[21] se desperdiciarán más recursos tratando de obtener y retener el monopolio, de modo que los costos sociales del monopolio podrían ser mayores. Tampoco generaría, en efecto, la discriminación perfecta de los precios el nivel de la producción competitiva. La discriminación de precios es costosa, en términos de la información y debido a la necesidad de impedir el arbitraje. Los costos en que se incurre al ejecutar un programa de discriminación de precios (parte de los costos de búsqueda de renta de la monopolización), bajo el supuesto plausible de que son costos variables, aumentarán el costo marginal de la empresa y, como sabemos, ello hará que el monopolista disminuya su producción. Además, la discriminación del precio distorsiona la competencia entre los clientes del monopolista al hacerlos pagar precios diferentes por insumos cuya producción cuesta lo mismo.

Al mismo tiempo, la discriminación del precio podría ser el único método de fijación de los precios que permite seguir operando a una empresa cuyo costo medio se encuentra en todos sus puntos por encima de su curva de demanda, aunque los consumidores en conjunto valúen el producto de la empresa en más que su costo total. El valor para los consumidores ocupa toda el área debajo de la curva de la demanda. A cualquier precio único por deba-

[20] Frederick M. Scherer y David Ross, *Industrial Market Structure and Economic Performance,* 494-496 (3ª ed., 1990); Jean Tirole, *The Theory of Industrial Organization,* 152-158 (1988).

[21] Adviértase que, en la gráfica IX.6, las ganancias totales del monopolio bajo la discriminación perfecta de los precios ocupan toda el área entre la curva del costo (*CMg* = *CM*) y la curva de la demanda.

jo de la intersección de la curva de la demanda con el eje vertical, algunos consumidores no pagarán el valor total que atribuyen al bien. Si ese excedente del consumidor (el valor menos el precio), sumado a los ingresos de la empresa, es igual o mayor a los costos de ésta, la empresa debe continuar operando, desde un punto de vista de eficiencia económica general; pero no lo hará a menos que los ingresos por sí solos cubran sus costos, y esto requerirá que discrimine precios en lugar de cobrar un precio único. La discriminación también puede aumentar la producción cuando las ganancias de la empresa están restringidas al nivel competitivo por la regulación; en ese caso, la discriminación del precio, llamada en este contexto "precios de Ramsey", generará una producción mayor que la del precio uniforme. Discutiremos más adelante, en XIII.5, los precios de Ramsey.

IX.6. Dos impedimentos a la monopolización (además de la entrada): la competencia por el mercado y la durabilidad

Si un comprador acepta tomar todos sus requerimientos de un bien particular de un solo vendedor, el vendedor adquirirá así un monopolio —en cierto sentido— sobre ese abasto. Pero el comprador no pagará más que el precio competitivo si se satisfacen tres condiciones: que haya varios vendedores que puedan abastecerlo, que estos vendedores no se coludan, y que no sea prohibitivo el costo de la celebración de un contrato eficaz para el periodo del monopolio. Porque entonces los vendedores lucharán entre sí para ofrecer un contrato atractivo, y el precio del contrato será el precio competitivo.[22] Pero supongamos que, una vez que el vendedor ha obtenido el contrato y ha empezado a cumplirlo, se requerirá largo tiempo para que otro vendedor pueda sustituirlo como proveedor del comprador (véase más adelante XIII.1 sobre por qué podría ocurrir esto). En tal situación, el comprador necesitará un contrato de largo plazo para protegerse contra el precio de monopolio, y los contratos de largo plazo podrían implicar costosas inflexibilidades (véase más adelante XIV.11).

Hay una dificultad especial en la monopolización de bienes duraderos,[23] como en el caso de la tierra. Si una persona poseyera toda la tierra de los Estados Unidos y deseara venderla a un precio o precios que maximizaran sus ganancias, como cualquier otro monopolista fijaría un precio al que se compraría sólo una porción de la tierra (recordemos que el monopolista siempre vende en la porción elástica de su curva de demanda, de modo que mediante el cobro de un precio más bajo vendería más de su producto).

[22] Harold Demsetz, "Why Regulate Utilities?", 11 *J. Law & Econ.,* 55 (1968); Yoram Barzel, "The Capture of Wealth by Monopolists and the Protection of Property Rights", 14 *Int'l Rev. Law & Econ.,* 393 (1994); *cf.* United States *vs.* El Paso Natural Gas Co., 376 U.S., 651 (1964).

[23] Ronald H. Coase, "Durability and Monopoly", 15 *J. Law & Econ.,* 143 (1972).

Pero, una vez que se completara la venta, el propietario tendría un incentivo para empezar a vender las porciones de la tierra restantes a un precio menor, hasta que finalmente se vendiera toda la tierra. Sabiendo esto, la gente no pagaría el precio inicial del propietario, y fracasaría el intento de monopolización.

Los diamantes son también un bien duradero, y, con todo, la producción de diamantes ha sido controlada desde hace mucho tiempo por un cártel internacional que limita la producción y cobra precios monopólicos.[24] Pero, al revés de la tierra, los diamantes se producen continuamente. Si el cártel de los diamantes aumentara la producción en un año y bajara considerablemente los precios a fin de alcanzar a un segmento de la comunidad que de ordinario no puede pagar por los diamantes, bajaría el valor de todos los diamantes que existen. Habiendo demostrado así su falta de confianza, el cártel no podría volver a cobrar jamás un precio tan alto como el que cobraba antes de su exagerada expansión. Sabiendo esto, los compradores de diamantes están dispuestos a pagar un precio de monopolio.

Incluso el terrateniente monopolista podría limitar creíblemente su producción… si rentara su tierra en lugar de venderla. Si después de haber rentado una parte de su tierra empezara a rentar o vender terrenos adicionales a precios menores, no podría cobrar un precio de monopolio tan elevado cuando llegara el momento de renovar el contrato de arrendamiento. Como sabían que el monopolista tenía un incentivo para respetar su limitación inicial de la producción, los compradores tendrían menos temor de pagar el precio monopólico completo por la renta.[25]

Así como el monopolio no es una condición suficiente del precio de monopolio, tampoco es una condición necesaria. Imaginemos un mercado de 100 vendedores, cada uno de los cuales produce 1 000 unidades, y supongamos que la expansión de la producción, ya sea por parte de las empresas existentes o por la entrada de empresas nuevas, es imposible. Cada vendedor tendrá el poder de monopolio: el poder de elevar el precio de mercado por encima del nivel competitivo. Por ejemplo, si uno de los 100 vendedores redujera su producción de 1 000 unidades a 900, la producción total del mercado caería de 100 000 a 99 900 unidades, y el precio de mercado se elevaría, al igual que si el monopolista del producto hubiese decidido reducir su producción en la misma cantidad.

[24] Godehard Lenzen, *The History of Diamond Production and the Diamond Trade,* 183-196 (F. Bradley [trad. al inglés], 1970).

[25] La seguridad sería mayor si el ingreso obtenido de los arrendamientos por el monopolista se fijara igual a un porcentaje de los ingresos del arrendatario. De este modo, toda acción del monopolista que redujera los ingresos del arrendatario reduciría sus propios ingresos directamente.

Lecturas recomendadas

Carlton, Dennis W., y Jeffrey M. Perloff, *Modern Industrial Organization*, caps. 4, 9, 16 (4a ed., 2005).
Hirshleifer, Jack, *Price Theory and Applications*, cap. 8 (5a ed., 1992).
Kaplow, Louis, *Competition Policy and Price Fixing* (2013).
Landes, William M., "Appendix: An Introduction to the Economics of Antitrust", en Richard A. Posner y Frank H. Easterbrook, *Antitrust: Cases, Economic Notes, and Other Materials,* 1055 (2a ed., 1981).
Posner, Richard A., "The Social Costs of Monopoly and Regulation", 83 *J. Pol. Econ.,* 807 (1975).
————, *Antitrust Law* (2a ed., 2001).
Scherer, Frederick M., y David Ross, *Industrial Market Structure and Economic Performance* (3a ed., 1990).

Problemas

1. ¿Por qué se encontrará siempre el precio de monopolio óptimo en la porción elástica de la curva de demanda del monopolista, a menos que el monopolista tenga costos marginales iguales a cero (¿cuál es la elasticidad de la demanda al precio que tal monopolista fijará?) o practique la discriminación del precio?

2. ¿Sería contribuyentemente negligente un consumidor en un sentido económico si, sabiendo que el precio de algún producto era un precio monopólico, no buscara asiduamente productos sustitutos?

3. ¿Dónde hemos visto antes el excedente del consumidor?

X. LAS LEYES ANTIMONOPÓLICAS[1]

X.1. Los cárteles y la Ley Sherman

Un contrato entre vendedores rivales para fijar el precio del producto que venden (o para limitar su producción, que es lo mismo) es como cualquier otro contrato en el sentido de que las partes no lo firmarían de no esperar que mejorara la situación de todas ellas. Pero perjudica a otros, los consumidores, quienes no son partes del contrato; y, como vimos en el capítulo anterior, cuando se toman en cuenta los efectos de sustitución y la tendencia de las ganancias monopólicas a transformarse en costos, los costos sociales exceden a las ganancias de los cartelistas. Es cierto que cada uno de los consumidores se encuentra en una relación contractual directa o indirecta con uno o más de los vendedores del cártel, de modo que podría parecer que el Teorema de Coase entrara en operación y que los consumidores pagarían a los vendedores para que expandieran su producción hasta el nivel competitivo. Sin embargo, esto implicaría que el estado final del mercado sería una discriminación perfecta del precio; y, aun si pudiera evitarse el arbitraje, serían prohibitivos los costos de la negociación por separado con cada consumidor por cada unidad. Éste es un ejemplo de la forma en que los costos de transacción pueden ser muy elevados incluso para las partes de una relación contractual (otro ejemplo es el monopolio bilateral). Por lo tanto, no es sorprendente que a fines del siglo XIX los tribunales estadunidenses se negaran a hacer cumplir los acuerdos de cárteles, alegando que iban contra la política pública.[2]

Eso no habría resuelto el problema, y no sólo porque la cooperación para el beneficio mutuo es posible sin la exigibilidad legal.[3] Además, la misma ines-

[1] Véanse Richard A. Posner y Frank H. Easterbrook, *Antitrust: Cases, Economic Notes, and Other Materials* (2ª ed., 1981); Richard A. Posner, *Antitrust Law* (2ª ed., 2001); Herbert Ovenkamp, *Federal Antitrust Policy: The Law of Competition and Its Practice* (3ª ed., 2005).

[2] La hostilidad hacia el monopolio, e incluso la valoración de la naturaleza y los beneficios sociales de los precios competitivos, son mucho más antiguos. Véase *Business, Banking and Economic Thought in Late Medieval and Early Modern Europe: Selected Studies of Raymond de Roover*, parte 4 (Julius Kirshner [comp.], 1974). El "precio justo" medieval ha sido ampliamente malinterpretado como un rechazo de los valores de mercado. Por el contrario, el precio justo era el precio de mercado cuando el mercado era competitivo. Raymond de Roover, "The Concept of the Just Price: Theory and Economic Policy", 18 *J. Econ. Hist.* 418 (1958); Diana Wood, *Medieval Economic Thought* (2002).

[3] Véase Ananth N. Madhavan, Robert T. Masson y William H. Lesser, "Cooperation for Monopolization? An Empirical Analysis of Cartelization", 76 *Rev. Econ. And Stat.*, 161 (1994). Véanse también IV.1 y VIII.7 anteriores.

tabilidad de un acuerdo de cártel imposible de hacer cumplir crea un incentivo para que los miembros del cártel se consoliden en una sola empresa; entonces el precio de monopolio se puede imponer sin depender de contratos.

Debemos hacer énfasis en la inestabilidad de un acuerdo de cártel imposible de hacer cumplir. El "producto" que "compra" una de las partes de tal acuerdo —la abstención de la competencia de precios— resulta difícil de inspeccionar. Si una de las partes pierde ventas, ello podría explicarse por muchas razones. Una es que uno de los competidores venda más barato que él. ¿Pero cómo podrá saberlo? Podría preguntar a los compradores que ha perdido, pero no podría confiar en sus respuestas. Podrían decirle que su competidor les estaba vendiendo más barato aunque no fuese cierto, a fin de inducirlo a disminuir su precio. Además, el competidor podría haberse adherido al precio del cártel pero mejorando su producto, lo que es un método particularmente sutil del engaño.[4]

Aunque el engaño implica necesariamente la venta a un precio menor (ajustado por la calidad), puede resultar rentable porque, dada la pendiente negativa en la curva de demanda, cualquier aumento en la producción total debe tener un precio por debajo del precio actual a fin de permanecer en la curva de demanda. Si es pequeña la expansión total de la producción del mercado generada por el engaño, la disminución de las ganancias del defraudador resultante de la disminución del precio de mercado podría ser menor que el aumento de sus ganancias por vender unidades adicionales. Sin embargo, si otras empresas pequeñas siguen el ejemplo del defraudador, el impacto acumulado sobre el precio del mercado y las ventas de los rivales, y, por lo tanto, sobre las ganancias de las empresas más grandes, los obligará a bajar sus precios defensivamente, y al final el precio del mercado bajará hasta el nivel competitivo… o por debajo de éste. Pues si las empresas grandes responden al engaño sólo mediante la disminución de su precio hasta el nivel del defraudador, éste obtendrá ganancias súper normales hasta entonces y ganancias normales cuando el precio vuelva al nivel competitivo; la rentabilidad del engaño se reduce pero no se elimina. A fin de dar una lección al defraudador, las empresas grandes podrían responder al defraudador bajando su precio más allá de su costo, imponiendo así pérdidas al defraudador como una sanción por su engaño. Por supuesto, las empresas grandes, los "aplicadores", perderán también. Pospondremos el tema de si las ganancias netas esperadas serán positivas hasta la discusión de la economía de los precios depredadores (véase después x.9).

[4] Por lo que toca a los problemas del cumplimiento forzoso de los cárteles, véanse George J. Stigler, "A Theory of Oligopoly", en *The Organization of Industry,* 39 (1968); John S. Mcgee, "Ocean Freight Rate Conferences and the American Merchant Marine", 27 *U. Chi. L. Rev.,* 191 (1960). Hay una abundante literatura empírica sobre los cárteles. Véanse las referencias en Richard A. Posner, *Antitrust Law,* 61 n. 13 (2ª ed., 2001), y un resumen útil en Dennis W. Carlton y Jeffrey M. Perloff, *Modern Industrial Organization,* cap. 5 (4ª ed., 2005).

Es posible que no se requiera vender por debajo del costo para disuadir a los desertores.[5] Si los participantes en el cártel se ponen de acuerdo sobre un precio detonador (un precio del mercado suficientemente por debajo del precio de cártel para probar que está ocurriendo un engaño) que será la señal para que los miembros del cártel vuelvan a la producción existente previamente en éste durante un periodo de tiempo especificado, la organización podría sobrevivir a los episodios de engaño. El arreglo del precio detonador cerrará el paso a la mayor parte de los engaños, aunque no a todos, y, cuando ocurra el engaño, el límite de tiempo para el periodo de castigo (el retorno a la producción existente antes del cártel y, por lo tanto, a la interrupción de precios y beneficios del cártel) podría permitir que el cártel continúe a pesar de la interrupción causada por el episodio de engaño.

La Ley Sherman (1890) imponía sanciones penales y civiles para los contratos y otras combinaciones, en restricción del comercio, la monopolización y las conspiraciones e intentos de monopolización. Las primeras decisiones interpretaban que la ley prohibía los cárteles; más tarde también se abolieron las fusiones de monopolio. Aunque las sanciones de la violación eran muy débiles al principio, aun las sanciones son razonablemente eficaces para impedir que los participantes en el cártel empleen ciertos instrumentos muy eficaces, pero también muy visibles, para la eliminación del fraude de sus miembros.[6] Tenemos un ejemplo en la agencia común de ventas; todas las ventas del producto del cártel se canalizan a través de la agencia, la que fija un precio uniforme, eliminando así toda competencia de precios entre los miembros del cártel.

Pero la Ley Sherman tuvo otro efecto sobre la cartelización que pudo haber sido ineficiente. La producción de un mercado monopolizado es menor que la de un mercado competitivo. Por lo tanto, cuando se forma un cártel, gran parte de la capacidad productiva del mercado se vuelve excesiva y debe ser retirada para economizar recursos. Pero los miembros del cártel, al estar preocupados por la posibilidad de que el cártel sea efímero, se mostrarán reacios a retirar capacidad porque podrían verse incapacitados para expandir rápidamente la producción si el cártel se derrumba. La agencia común de ventas y otros métodos eficientes para el cártel suprimidos por la Ley Sherman facilitan el retiro de la capacidad excedente al aumentar la estabilidad y la longevidad del cártel que los emplea, y en este sentido podrían ser menos dispendiosos de recursos que el cártel oculto que los ha remplazado. Pero, contrario a esta conclusión, está el hecho de que, en la medida en que el precio promedio del cártel "eficiente" sea mayor debido a que el cártel es más

[5] Véase Glenn Ellison, "Theories of Cartel Stability and the Joint Executive Committee", 25 *RAND J. Econ.,* 37 (1994); Jonathan B. Baker, "Identifying Cartel Policing under Uncertainty: The U.S. Steel Industry, 1933-1939", 32 *J. Law & Econ.,* S47 (1989).

[6] George J. Stigler, "The Economic Effects of the Antitrust Laws", en *The Organization of Industry*, nota 4 anterior, en 259.

estable y dura más, los efectos de una mala asignación del cártel eficiente serán mayores, lo que es muy probable que contrarreste los ahorros en costos derivados de eludir el exceso de capacidad.

En la aplicación de la Ley Sherman contra cárteles y conspiraciones se ha hecho hincapié en la prueba de que existe un acuerdo para la fijación de precios —una cuestión legal—, antes que en la prueba de los efectos de la conducta de los vendedores sobre el precio y la producción —una cuestión económica—. Una consecuencia no buscada de este hincapié es que los cárteles que tienen mayores probabilidades de ser descubiertos y enjuiciados son aquellos donde los efectos de precio y producción son pequeños. Tales son los cárteles: *1)* de muchos miembros, de modo que es más probable que uno de ellos se disguste e informe acerca de los otros; *2)* que dependen de la negociación y el acuerdo explícitos y reiterados, lo que provee la prueba esencial de la violación, y *3)* que tienden a llenarse de engaños y a derrumbarse pronto en medio de recriminaciones mutuas: circunstancias que crean oportunidades para que el gobierno o los responsables privados del cumplimiento consigan testigos dispuestos a ofrecer pruebas del acuerdo. Es menos probable que el cártel de funcionamiento adecuado genere pruebas del acuerdo efectivo. Lo que la ley castiga sobre todo es el intento de fijación de precios. La conspiración consumada escapa a menudo a la atención.[7]

El análisis económico identifica las varias características que indican la predisposición de un mercado a la fijación efectiva de los precios sin requerir muchos (o quizás incluso ningunos) acuerdos públicos o comunicación entre las empresas:[8]

1. *El número de (grandes) vendedores*. Entre menor sea ese número, menores serán los costos de la coordinación de sus actividades.

2. *No hay una franja de vendedores pequeños*. Hay una diferencia incluso en un mercado altamente concentrado, como aquel en el que la participación de mercado agregada de las cuatro empresas más grandes es de 80%, tanto si hay sólo otra empresa como otras 10 empresas. La coordinación de la fijación de precios de 5 empresas es más fácil que la de 14, y aun cuando puede resultar innecesario obtener el acuerdo de las empresas pequeñas a fin de coludirse eficazmente —éstas quizá no puedan expandir su produc-

[7] Howard P. Marvel, Jeffrey M. Netter y Anthony M. Robinson, "Price Fixing and Civil Damages: An Economic Analysis", 40 *Stan. L. Rev.,* 561 (1988), descubren que, en casi la mitad de los casos penales de antimonopolio del Departamento de Justicia, un fallo a favor del gobierno *no* fue seguido por un juicio de daños, aunque tal juicio de seguimiento habría sido fácil de litigar para obtener un fallo o un arreglo de pago triple de daños. ¿Qué implica esto acerca de la eficiencia de la política de cumplimiento legal del Departamento? Véanse también Peter Asch y J. J. Seneca, "Is Collusion Profitable?", 58 *Rev. Econ. & Stat.,* 1, 8 (1976); David F. Lean, Jonathan D. Ogur y Robert P. Rogers, "Does Collusion Pay… Does Antitrust Work?", 51 *S. J. Econ.,* 828, 839 (1985); Michael F. Sprout, "Antitrust and Prices", 101 *J. Pol. Econ.,* 741 (1993).

[8] Para una lista más larga y una discusión más completa, véase Posner, nota 4 anterior, en 69-79.

ción con rapidez suficiente para contrarrestar la restricción de la producción de los vendedores que se coluden a corto plazo, y en todo caso podrían tener costos mayores que los de las empresas que se coluden—, cualquier parte del mercado que se encuentre fuera del círculo que se colude constituye una limitación al poder de los vendedores coludidos para elevar el precio del mercado. Y lo importante no es la participación de mercado de la empresa individual al margen, sino la del margen en conjunto. Si las empresas que agregan 20% del mercado, aun siendo numerosas, pueden expandir su producción 25%, esto limitará el precio cobrado por el cártel y, por ende, sus ganancias y, en consecuencia, la probabilidad de que un cártel se forme o persista.

Esta observación se relaciona con la elección de métodos de medición de la concentración. El más común de tales métodos es la participación de mercado combinada de las cuatro u ocho empresas más grandes. Pero un índice mejor, utilizado ahora por las autoridades encargadas del cumplimiento y que empieza a filtrarse también al derecho casuístico, es el índice de Herfindahl (oficialmente llamado Índice de Herfindahl-Hirschman, o IHH), porque toma en cuenta a la franja de empresas que permanecen al margen. El Herfindahl de un mercado es la suma del cuadrado de la participación de cada empresa en el mercado multiplicada por 1 000. Si cada una de las cuatro empresas más grandes tiene 20% del mercado y cada uno de una franja de 10 vendedores pequeños tiene 2%, el índice de Herfindahl sería 1 640. Pero si en lugar de una franja de 10 vendedores pequeños hubiese otra empresa con 20%, el Herfindahl sería 2 000, indicando correctamente que este mercado estaba más concentrado. Aun así, la razón de concentración de cuatro empresas sería la misma en los dos mercados (80 por ciento).

La diferencia entre una franja de firmas con un total de 20% del mercado y una firma con 20% es que la franja es un contrapeso al poder de mercado de las cuatro firmas más grandes, mientras que una firma con un 20% es candidata para sumarse al cártel y, por esa vía, permitirle controlar el 100% del mercado.

3. *La homogeneidad del producto*. Entre más homogéneo sea un producto, más difícil será el engaño mediante la alteración de la calidad del producto; el cambio será evidente.

4. *La elasticidad de la demanda respecto del precio*. En igualdad de las demás circunstancias, entre menos elástica sea la demanda al precio competitivo, mayores serán las ganancias que generará un precio monopólico y en consecuencia será mayor el incentivo para la monopolización. Intuitivamente, entre menor sea la disminución de la cantidad demandada en respuesta a los aumentos de precio, mayor libertad tendrá el monopolista para aumentar el precio.

5. *La condición de la entrada*. Si la entrada puede realizarse con rapidez y quienes ingresan no tienen costos a largo plazo mayores que los de los miembros del cártel, las ganancias de la cartelización serán pequeñas, al igual que

el incentivo para cartelizarse. Esto se relaciona con el punto anterior: Mientras la entrada sea más fácil, mayor será la elasticidad de la demanda; los aumentos de precios derivarán en grandes aumentos de cantidad a medida que entren participantes nuevos, y, por lo tanto, las empresas que lleven a cabo el aumento de precios perderán ventas rápidamente.

6. *¿Está el mercado creciendo, declinando o permaneciendo constante a lo largo del tiempo?* Si la demanda está creciendo, la cartelización será difícil de hacer respetar, porque un vendedor cuya participación en el mercado está disminuyendo tendrá dificultades para determinar si simplemente no le está yendo tan bien como a sus rivales en la atracción de nuevos compradores al mercado, o si sus rivales están bajando el precio. (Pero, por otra parte, si los cartelistas, en lugar de ponerse de acuerdo sobre el precio, convienen en la tasa de construcción de nuevas plantas, el cumplimiento será más fácil de vigilar.) Si la demanda es constante o declinante, es más probable que una pérdida de mercado se deba (y se atribuya) al engaño sobre el precio del cártel. Un mercado declinante es especialmente favorable para la cartelización. El riesgo de la quiebra será mayor porque los costos fijos no pueden disminuirse reduciendo la producción; por lo tanto, la competencia de precios parecerá peculiarmente destructiva (véase después XIII.8). Más importante aún, la entrada no es una amenaza seria para un cártel en un mercado declinante.

7. *Una razón elevada de los costos fijos a los variables*. En un mercado donde los costos fijos son una proporción elevada de los costos totales, la competencia puede ser inestable y conducir a quiebras frecuentes. Dado que por definición los costos fijos no pueden reducirse para ajustarse a una disminución de la demanda, una empresa que tenga grandes costos fijos es muy vulnerable a una recesión económica o a cualquier otro desarrollo inesperado que pueda obligarla a disminuir su producción o bajar su precio a fin de sobrevivir; sus ingresos estarán bajando más deprisa que sus costos. (Este punto tiene una relación estrecha con el anterior; si una empresa tiene una deuda fija, una disminución en la demanda, al reducir la producción de la firma, incrementará su razón de los costos fijos a los costos variables.) Si la quiebra es más costosa para los administradores de una empresa de lo que sería el mero fracaso para obtener la diferencia existente entre una tasa de rendimiento a nivel competitivo y la de nivel monopólico,[9] considerarán los beneficios derivados del

[9] Los administradores podrían tener un capital humano específico de la empresa, es decir, habilidades especializadas para la empresa, de modo que si pierden su empleo tendrán dificultades para encontrar otro empleo con un sueldo comparable. (Ellos también podrían tener una participación desproporcionada de su riqueza atada a la empresa en virtud de haber recibido opciones sobre acciones.) Es improbable que los accionistas que no forman parte de la administración o gerencia tengan la misma preocupación. Es mucho más fácil diversificar nuestro capital financiero que nuestro capital humano. Por lo tanto, los accionistas pueden desear que sus administradores asuman más riesgos que los que éstos desean asumir… y, si los administradores son agentes imperfectos de los accionistas, es posible que no lo hagan. Tratan de minimizar el riesgo de la quiebra a toda costa, y un procedimiento para hacerlo es coludiéndose con sus competidores.

precio monopólico más importantes de lo que serían en una industria donde el precio competitivo no pondría en peligro la solvencia de la empresa.

8. *Los precios pueden cambiarse rápidamente*. Entre más aprisa pueda reaccionar un vendedor ante la reducción del precio de un competidor, menos rentable será la reducción del precio porque éste será igualado antes de que la empresa que redujo el precio pueda obtener gran ventaja de ello. Las políticas deliberadas de igualación de los precios pertenecen a la estrategia, no a la estructura o los antecedentes, y las discutiremos más adelante. Pero a menudo la naturaleza del producto o de su sistema de distribución determina la frecuencia con la que puede cambiarse el precio. Entre menos frecuentemente pueda cambiarse, más largo será el periodo en el que una empresa que baja los precios puede esperar beneficiarse de su precio menor.

9. *La estructura del lado comprador del mercado*. Si hay muchos compradores de igual tamaño, el engaño al cártel requerirá muchas transacciones y, por lo tanto, será grande la probabilidad de detección por parte de los otros miembros del cártel. Pero si hay pocos compradores (grandes), un miembro del cártel podría engañar provechosamente simplemente atrayendo a uno o dos clientes de otro miembro del cártel. Pocas transacciones serán necesarias, y esto disminuirá la probabilidad de la detección. De igual modo, podría resultar difícil para la víctima del engaño determinar si perdió ventas debido a la reducción del precio o a factores aleatorios; los pocos compradores que ha perdido serán una muestra demasiado pequeña para extraer una deducción confiable en cuanto a la razón por la cual desertaron.

El análisis económico puede indicar también cuál tipo de prueba revela que un mercado *está* siendo cartelizado, a diferencia de que sólo exista la probabilidad de que se cartelice.

1. Un ejemplo es el de la discriminación de precios en todo el mercado: un método de explotar el poder monopólico,[10] como hemos visto. Pero debemos distinguir entre la discriminación persistente o sistemática, que es un síntoma del monopolio, y la discriminación temporal o esporádica que acompaña de ordinario al movimiento de un equilibrio a otro en un mercado competitivo, y también de la desintegración o el debilitamiento de un cártel a resultas del engaño, que a menudo asumirá la forma de descuentos selectivos —es decir, discriminatorios— respecto al precio del cártel. Las cláusulas de nación más favorecida (véase después XXXI.2), debido a que pueden ser instrumentos eficaces para impedir la discriminación de los precios, pueden ser indicativas de una colusión tácita.[11] Dicha cláusula requiere que nuestro socio contractual

[10] "Todo el mercado" es una calificación importante. Una empresa individual que fabrica un producto para el cual no existe un sustituto cercano podría ser capaz de discriminar precios, debido a que tiene cierto poder monopólico; esto no significa que el mercado compuesto por esa empresa y sus competidores (imperfectos) esté cartelizado.

[11] Lo que no sugiere que éste sea el propósito o el efecto habitual de tales cláusulas; a menudo tienen el propósito inocente de introducir la flexibilidad de los precios en contratos a largo

nos trate tan bien como trata a nuestros competidores. Si los pequeños compradores de un mercado, temiendo que sus competidores grandes puedan obtener una ventaja sobre ellos al conseguir descuentos secretos de los miembros de un cártel de vendedores, solicitan la inclusión de una cláusula de nación más favorecida en sus contratos con los vendedores, estos últimos podrían acceder con gusto, porque el efecto de tal concesión sería impedir que el cártel se erosione por causa de tales descuentos. Los compradores que consideren ejercer presión a favor de estas cláusulas o las consientan tendrán que tomar una decisión difícil entre el fortalecimiento del cártel y la pérdida a manos de sus competidores, pero es posible que el equilibrio favorezca al primer curso de acción. ¡Es posible que los vendedores incluso les paguen para que acepten tales cláusulas!

La discriminación que es un síntoma de la cartelización debe distinguirse también de la discriminación que se debe al poder de mercado de empresas individuales. Cuando los productos de las empresas competidoras no son idénticos, aunque sean suficientemente similares para convertirse en sustitutos cercanos, cada empresa podría tener cierto poder monopólico unilateral. Por ejemplo, la diferencia tradicional del precio entre las películas cinematográficas de estreno y las de repetición es discriminatoria porque el costo de la distribución no es mayor para la segunda corrida que para la exhibición de estreno, pero la discriminación se facilita merced al poder monopólico individual de los productores de las películas, derivado de sus *copyrights*, antes que a una colusión entre los productores. Un ejemplo similar es el de la diferencia de precio entre las versiones de pasta dura y de pasta blanda del mismo libro; invariablemente, la diferencia supera a la diferencia del costo, aunque la industria editorial sea muy competitiva. Otro ejemplo es el precio mayor al que se venden las revistas académicas a las bibliotecas que a los individuos; otro más, los descuentos de fines de semana que ofrecen las aerolíneas que discriminan contra sus clientes comerciales (de baja elasticidad).

Todos los ejemplos anteriores, menos el último, involucraban a la propiedad intelectual. No sólo se encuentra tal propiedad en el otro extremo del espectro de la diferenciación del producto en relación con los bienes remplazables y muebles sino que, en virtud de que el costo marginal de la propiedad intelectual tiende a ser menor que su costo total medio (el ejemplo más claro es el del software de computadora, cuya creación puede resultar inmensamente costosa pero cuyas copias pueden hacerse y distribuirse virtualmente sin costo),[12] la discriminación de precios es una estrategia atractiva

plazo. Un comprador podría mostrarse reticente a firmar un contrato a largo plazo con un precio fijo sin la protección que le otorga una cláusula de nación más favorecida, la que ajustaría el precio del contrato a las condiciones cambiantes del mercado. Véase Keith J. Crocker y Thomas P. Lyon, "What Do 'Facilitating Practices' Facilitate? An Empirical Investigation of Most-Favored Nation Clauses in Natural Gas Contracts", 37 *J. Law & Econ.*, 297 (1994).

[12] Supongamos que el costo de creación del software, un costo fijo, es de 3 000 dólares, pero

para aumentar la producción al mismo tiempo que se cubren los costos totales. Con todo, el hecho mismo de que estos bienes no sean remplazables y muebles implica que los descuentos *uniformes* entre diversas marcas podrían resultar sospechosos de ser el producto de la colusión; por ejemplo, si todas las editoriales académicas cobraran a las bibliotecas cinco veces más que a los individuos.

Surge otra dificultad para caracterizar adecuadamente las diferencias del precio del hecho de que, cuando dos productos tienen costos conjuntos o comunes (el cuero y la carne de un novillo es el ejemplo tradicional), los precios para clientes diferentes pueden ser diferentemente proporcionales a los costos marginales del servicio de cada cliente sin ser discriminantes en ningún sentido relevante para la inferencia de un poder monopólico. El cuero y la carne no reciben el mismo precio, aunque el costo marginal (el costo de la crianza del novillo) sea el mismo (antes de ser procesados). Pero esto no es discriminación.

2. Otro síntoma de la cartelización, ilustrado por el cártel internacional del petróleo, la Organización de Países Exportadores de Petróleo (OPEP), es una disminución, a lo largo del tiempo, de la participación de mercado de las empresas más grandes en el mismo, lo que podría indicar que tales empresas han venido cobrando un precio monopólico que ha atraído a nuevos miembros, quienes les han arrebatado clientes cobrando precios menores.[13]

3. Un tercer tipo de prueba es el mantenimiento del precio de reventa a nivel de toda la industria, lo que, a menos que se justifique por las razones que discutiremos enseguida, puede haberse adoptado para impedir que los vendedores engañen dentro de su cártel al vender a los distribuidores con un margen de ganancia reducido.[14] Contrario al precio de venta al menudeo, el precio al que se venden los bienes al distribuidor no es visible para los competidores del vendedor, por lo que reducir ese precio es un método de engaño efectivo.

4. Otro síntoma de cartelización son las participaciones del mercado demasiado estables para ser un resultado plausible de la actividad competitiva normal entre los vendedores.

que el costo de la elaboración y venta de las copias es de sólo 1 dólar por copia. Entonces, a un nivel de producción de 1 000 copias, el costo marginal es un dólar, pero el costo total medio ([3 000 dólares + 1 000 dólares] ÷ 1 000) es 4 dólares.

[13] Robert S. Pindyck, "Gains to Producers from the Cartelization of Exhaustible Resources", 60 *Rev. Econ. & Stat.,* 238 (1978); Dermot Gately, "A Ten-Year Retrospective: OPEC and the World Oil Market", 22 *J. Econ. Lit*., 1100 (1984); James M. Griffin y William S. Neilson, "The 1985-1986 Oil Price Collapse and Afterwards: What Does Game Theory Add?", 32 *Econ. Inquiry*, 543 (1994); S. Gürcan Güllen, "Is OPEC a Cartel? Evidence from Cointegration and Causality Tests", 17 *Energy J*., 43 (1996).

[14] Si un distribuidor baja su precio, los competidores del proveedor de ese distribuidor pueden sentirse inseguros acerca de que la rebaja del precio refleje un cambio en los costos de operación del minorista o una rebaja del precio del proveedor; el mantenimiento del precio de reventa elimina esta incertidumbre.

5. Otra prueba es el hecho de que las variaciones regionales del precio no puedan explicarse por las diferencias regionales del costo o de la demanda. La implicación es que los precios más altos son supracompetitivos, mientras que es de suponer que los más bajos cubren los costos de los vendedores que cobran precios menores.

6. Una elevada elasticidad de la demanda al precio de mercado corriente, aunada a una ausencia de buenos sustitutos del producto (es decir, comparables en costo y valor con el producto para el que son sustitutos), sugiere que la elasticidad elevada es el resultado del precio monopólico; pues recordemos que un monopolista seguirá aumentando el precio hasta que esté vendiendo en la proporción elástica de su curva de demanda. Éste es un ejemplo de la diferencia existente entre las condiciones económicas que facilitan la cartelización y las condiciones económicas que la ponen de manifiesto. Si la demanda es inelástica al precio competitivo, esto vuelve atractiva a la cartelización porque un aumento del precio no causará una disminución proporcional de la cantidad demandada. Pero si se forma un cártel en consecuencia, éste continuará aumentando el precio hasta que se encuentre dentro de la región elástica de la curva de demanda del mercado. La ausencia de bienes sustitutos es relevante debido a que la elasticidad elevada de la demanda del producto de una empresa implica, frente a sustitutos cercanos, competencia efectiva y no precios de monopolio: si el precio del producto se eleva aunque sea un poco, los compradores lo cambiarán por los sustitutos.

7. En ocasiones, quizá pueda inferirse la colusión de la presencia o el patrón de tasas de rendimiento anormalmente elevadas. Pero el problema de medición es agudo. Las tasas extraordinariamente elevadas podrían deberse a la eficiencia superior y no a la colusión, o a un crecimiento extremadamente rápido en la demanda (lo que lleva a los vendedores a racionar el suministro elevando el precio), a convenios específicos y otros factores inocentes. Sin embargo, lo que es igualmente importante, la ausencia de ganancias anormalmente elevadas no implica necesariamente la ausencia de colusión. Es posible que los miembros del cártel hayan visto desaparecer sus ganancias monopólicas por las diversas formas de la competencia en variables distintas del precio. En efecto, en una industria declinante, que como sabemos es un contexto particularmente favorable para la colusión, es improbable que las empresas sean rentables, por lo menos en un sentido contable (una calificación importante, ya que esa empresa podría tener un flujo de efectivo considerable si no tratara de remplazar la planta a medida que ésta se deprecia y, aun así, todavía puede asumir la depreciación de la planta reduciendo sus ganancias reportadas).

Cuando los "estudios de evento" de los precios de las acciones demuestran un aumento inexplicable del valor de los vendedores sospechosos en cuestión, esto podría evidenciar la colusión.[15] En la formación de un cártel, por

[15] Los estudios de eventos utilizan métodos estadísticos para aislar el efecto de un evento

ejemplo, las ganancias de las empresas más pequeñas del mercado aumentarán en una proporción mayor que las ganancias de las empresas más grandes. La razón es que las empresas más pequeñas tienden a permanecer fuera del cártel (no conviene dejarlas que se unan) y, por lo tanto, a producir los mismos niveles que antes, o mayores aún, mientras que las empresas más grandes están disminuyendo su producción a fin de aumentar el precio de mercado. Las empresas más pequeñas serán entonces capaces de gorronear y aprovecharse gratuitamente del precio del cártel elevando sus precios a ese nivel. De modo que la estrategia óptima para una empresa pequeña en un mercado cartelizado es ser un no miembro gorrón. Sin embargo, una empresa grande no puede gorronear porque la reducción de su precio por debajo del precio del cártel debilitaría el precio de mercado de forma significativa. Ser grande significa que una reducción en el precio expandirá su producción desde una base ya grande, con efectos dramáticos sobre la producción de las otras empresas.

A medida que pasa el tiempo, las ganancias de las empresas pequeñas deberán declinar mientras que el precio monopólico atrae al mercado a nuevas empresas (las que serán presumiblemente pequeñas, por lo menos al principio) que ganan sólo un rendimiento normal, además de inducir la expansión de las empresas pequeñas que ya se encuentran en él. Por lo tanto, la brecha de rentabilidad existente entre las empresas grandes y las pequeñas tenderá a cerrarse al mismo tiempo que las ganancias de ambas declinan porque las empresas grandes están destinando más recursos a la competencia en variables distintas del precio en el esfuerzo de obtener una parte mayor de las ganancias que el precio de cártel ha hecho posibles.

8. Por último, la colusión puede inferirse a veces de una correlación negativa entre el número de empresas en el mercado y el precio de mercado. La teoría de la competencia implica que el precio se determina por el costo; el número de empresas debiera ser irrelevante. Si, por el contrario, el precio es una función negativa del número de empresas, aumentando a medida que ese número disminuye y bajando a medida que ese número aumenta, esto implica colusión, porque la colusión es más eficaz entre menor sea el número de empresas en el mercado.

particular, tal como una fusión, un anuncio público de noticias buenas o malas o la formación o disolución de un cártel, sobre el precio de las acciones de una compañía. Un estudio de eventos compara el precio de las acciones antes y después del evento de interés, haciendo correcciones para circunstancias adicionales al evento que pudiesen haber influido en el cambio de precio. Véanse en general Sanjai Bhagat & Roberta Romano, "Event Studies and the Law: Part 1: Technique and Corporate Litigation", 4 *Am. L. & Econ. Rev.,* 141 (2002); *ibid.*, "Part II: Empirical Studies of Corporate Law", 4 *Am. Law & Econ. Rev.*, 380 (2002), y un estudio ilustrativo en William Knapp, "Event Analysis of Air Carrier Mergers and Acquisitions", 72 *Rev. Econ. & Stat.,* 703 (1990).

X.2. Colusión tácita

Supongamos que las pruebas económicas apuntaran inequívocamente hacia la colusión, pero que no existen pruebas de la comunicación efectiva entre los coludidos. Cada uno de ellos parecía haber decidido independientemente cuál sería el precio de la maximización conjunta y haberlo cobrado, asumiendo el riesgo de que otros vendedores engañaran o tuvieran una idea diferente de lo que era el precio de la maximización conjunta. Pero si hubiera muy pocos vendedores y un producto homogéneo, y se satisficieran también las otras condiciones facilitadoras enumeradas antes, la colusión tácita —la colusión lograda sin comunicación entre los participantes de la colusión— podría funcionar. ¿Debiera la ley abstenerse de actuar en ese caso, o debiera castigar o prohibir al cártel tácito como si hubiera un cártel expreso?[16] Los escépticos podrían preguntar: si se prohibiera la colusión tácita, ¿cómo fijaría su precio un vendedor de un mercado en el que las condiciones facilitadoras permitieran la convergencia en un precio de maximización conjunta? ¿Se vería obligado a adoptar un procedimiento de fijación del precio igual al costo más un margen de ganancia? ¿Se impediría la entrada en mercados cartelizados por el hecho de que un neófito que, tras haber logrado entrar en tal mercado y cobrar el precio prevaleciente en el mercado, pudiera ser enjuiciado como participante en una colusión tácita?

La cuestión del remedio, y una cierta imprecisión en el concepto mismo de la "colusión tácita" (una expresión que suena un poco como estar de acuerdo mientras no se está de acuerdo), se destacan en el ejemplo siguiente. Supongamos que los competidores en un mercado oligopólico no tratan de vender su producto a los "durmientes" del otro, siendo éstos clientes fieles de un vendedor o de otro, quienes no compran. El competidor X se dice a sí mismo: "no 'despertaré' a ninguno de los durmientes de Y porque, si lo hago, él despertará a los míos y los dos estaremos en peores condiciones". X y Y no se comunican entre sí, y X sólo está especulando que, si roba a los durmientes de X, Y robará a los suyos. ¿Es ésta una conspiración por la que se puede castigar a X y Y? Y, si se les castiga, ¿qué harán después? ¿Deben robar los durmientes del otro para evitar el castigo futuro? Una orden judicial para hacer tal cosa parecería muy extraña.

[16] Véase un "no" enfático para la primera pregunta y un "sí" para la segunda en Louis Kaplow, *Competition Policy and Price Fixing* (2013), amplificando y mejorando Richard A. Posner, *Antitrust Law*, cap. 3 (2001), y Posner, "Oligopoly and the Antitrust Laws: A Suggested Approach", 21 *Stan. L. Rev.,* 1562 (1969). Véanse también Williamson Oil Co. *vs.* Philip Morris USA, 346 F.3d, 1287, 1305-1310 (11th Cir., 2003); In re High Fructose Corn Syrup Antitrust Litigation, 295 F.3d, 651, 654 (7th Cir., 2002); JTC Petroleum Co. *vs.* Piasa Motor Fuels, Inc., 190 F.3d, 775, 780-781 (7th Cir., 1999); Wallace *vs.* Bank of Bartlett, 55 F.3d, 1166 (6th Cir., 1995); E.I. Du Pont de Nemours & Co. *vs.* FTC, 729 F.2d, 128 (2nd Cir., 1984).

O consideremos este caso común: X y Y son competidores. Y aumenta su precio. X hace lo mismo. ¿Es esto una conspiración? Quizás X piense que Y tiene conocimientos de la demanda del mercado de los que él, X, carece. Quizás X tenga miedo de que, aunque sus ventas aumenten si no hace lo mismo, las ganancias de Y aumenten aún más y, por lo tanto, los accionistas de X estén insatisfechos. Si estas reflexiones convencen a X —de nuevo sin ninguna comunicación con Y— de aumentar su precio, ¿es ésta una conspiración que amerita acción procesal? Una vez más tengo mis dudas.

Si se creyera que la prohibición de la colusión tácita es poco factible, ¿se podría obtener un alivio eficaz de otras maneras, por ejemplo, al ordenar un cambio en la administración de las empresas en connivencia, prohibir las cláusulas de naciones más favorecidas o prohibir la notificación de los cambios de precio por más de unos pocos días o semanas de antelación? En principio, sí, pero, aparte de prohibir las cláusulas de nación más favorecida, las formas de compensación sugeridas serían muy difíciles de administrar para un tribunal o de prever y evitar para los vendedores. Un vendedor en dicho mercado tendría que adoptar alguna forma de fijación del precio igual al costo más un margen de ganancia para mantenerse fuera de problemas, porque si considerara los factores de demanda en absoluto en la fijación de precios estaría considerando inevitablemente el efecto de su precio en los precios de sus competidores y, por lo tanto, en la demanda de su bien. Contra toda lógica, podría obstaculizarse la nueva entrada en los mercados oligopólicos porque el neófito se preocuparía de ser castigado por la colusión tácita si estableciera su precio justo por debajo del de las empresas ya establecidas con el fin de maximizar su ganancia derivada de la entrada.

Sin embargo, aun concediendo todas estas dudas, uno puede imaginar un caso en el que uno o más de los ocho tipos de pruebas de colusión puramente económicas mencionados anteriormente excluyeran alguna deducción sobre el comportamiento de las empresas en un mercado propicio a la colusión distinto del de la colusión tácita. Ésa podría ser una base aceptable para una declaración de ilegalidad[17] y para diferenciar mis dos ejemplos anteriores. Pero prevalecería el problema de lo que tendrían que hacer las empresas en el futuro para mantenerse fuera de problemas, de si incluso sólo pensar en la competencia para decidir qué precios cobrar se consideraría como prueba de colusión tácita.

Se me concede una pausa adicional por el hecho de que un caso de antimonopolio contra la colusión tácita dependería *por completo* de pruebas económicas; ya que existe el riesgo de que los tribunales manejen mal las pruebas económicas. Existe una larga historia de esto. Un ejemplo conocido es el caso del monopolio de la U.S. Steel.[18] Al fallar a favor de la parte demandada, la

[17] Como se discutió ampliamente en Louis Kaplow, *Competition Policy and Price Fixing*, especialmente en los caps. 10-11 (2013).

[18] United States *vs.* United States Steel Corp., 251 U.S., 417 (1920).

Corte Suprema se impresionó al saber que la participación de mercado de la U.S. Steel había bajado sostenidamente después de la combinación de siderúrgicas en competencia para formar la corporación (¡y que sus competidores no se habían quejado de sus tácticas competitivas!). La Corte no reconoció el comportamiento monopólico. (Éste es el ítem número 2 en mi catálogo de evidencias económicas de monopolio o de fijación de precio por los cárteles.) El establecimiento de un precio monopólico crea un incentivo para que accedan al mercado nuevos vendedores. El monopolista tiene tres opciones. Puede dejar de cobrar un precio monopólico, a fin de desalentar la entrada; puede no hacer nada; o puede disminuir su producción. El primer curso de acción derrota por completo el propósito del monopolio. Bajo el segundo o el tercero, el monopolista obtiene algunas ganancias monopólicas, por lo menos temporalmente, de modo que esperaríamos que siguiera cualquiera de estos dos cursos. Ambos conducen a una disminución de su participación en el mercado. El tercero podría ser superior, porque una reducción en la producción aumentaría el precio de mercado… pero también podría estimular la entrada adicional.

X.3. Problemas con las pruebas económicas

Utilizar la ley antimonopolio en contra de la colusión tácita requeriría de un grado de sofisticación económica del que es probable que carezcan los jueces y abogados. Existe una larga e indecorosa historia de desconcierto judicial en el tratamiento de las pruebas económicas en un caso de antimonopolio. En el segundo caso de American Tobacco,[19] la Corte Suprema, al sostener que los grandes fabricantes de cigarrillos habían conspirado para eliminar la competencia, pensó que era ominoso que los fabricantes hubiesen aumentado sus precios durante la Gran Depresión del decenio de 1930, a pesar de que se trataba de un periodo de costos declinantes al igual que la demanda. Sabemos del capítulo IX que un monopolista (o cártel) maximizador de sus ganancias disminuirá su precio cuando bajen sus costos, y que una disminución de la demanda lo hará disminuir su precio (si el costo marginal aumenta junto con el nivel de producción) o mantenerlo constante (si el costo marginal es constante). El monopolista aumentará su precio sólo si el costo marginal disminuye cuando la producción se encuentra en la región relevante, lo que en el caso del tabaco habría implicado, como veremos en el capítulo XIII, que la industria cigarrera tuviera un monopolio natural; esto es muy poco plausible. Otra posibilidad es que un cártel cigarrero se haya formado primero en la Depresión, pero no hay ninguna prueba de esto. Una más es que la demanda, además de bajar, se estuviese volviendo menos elástica. Eso alentaría al monopolista a aumentar el precio, y este efecto podría dominar su incentivo

[19] American Tobacco Co. *vs.* United States, 328 U.S., 781 (1946).

para disminuir el precio porque la demanda estaba declinando. Pero, durante las depresiones, las demandas se vuelven más elásticas, no menos, porque la gente compra entonces con mayor cuidado. Este presentimiento se ve fortalecido por las incursiones competitivas realizadas por las marcas tabacaleras menores cuando las compañías tabacaleras mayores aumentaron sus precios. Una última posibilidad es que, dado que, como sabemos, los cárteles tienden a ser más estables en los mercados que están a la baja que en los mercados en expansión, los fabricantes de cigarrillos que enfrentaron una caída en la demanda pensaran que podrían aumentar el precio sin incitar al engaño o atraer la entrada… aunque, si eso fue lo que pensaron, estaban equivocados, al menos respecto a la entrada.

Ignorar a la economía para favorecer las pruebas del acuerdo puede dar lugar a una observancia excesiva así como a un insuficiente cumplimiento de los principios de antimonopolio. Consideremos esta demanda reciente contra dos empresas canadienses filiales que producían ácido sulfúrico como un subproducto no deseado de sus operaciones mineras.[20] Dado que el mercado canadiense para el ácido sulfúrico es muy pequeño, las empresas comenzaron a buscar en el gran mercado estadunidense. Para acceder a éste necesitaban una red de distribución estadunidense. Los candidatos lógicos para constituir dicha red eran los productores estadunidenses de ácido sulfúrico, quienes manufacturaban el ácido a partir del azufre. Pero la venta de ese ácido sulfúrico (llamado "ácido virgen") no era muy rentable, por lo que las empresas canadienses vieron una oportunidad para convencer a los productores estadunidenses de distribuir su ácido sulfúrico ("ácido fundidor") en lugar de producir su propio ácido. Las empresas estadunidenses eran sus propias distribuidoras y, por lo tanto, tenían las relaciones comerciales necesarias para la distribución del ácido fundidor de las empresas canadienses en los Estados Unidos.

El esfuerzo de persuasión tuvo éxito. El ácido fundidor era más barato, y los productores estadunidenses temían que, a menos que aceptaran convertirse en distribuidores para las empresas canadienses, éstas crearían su propia red de distribución en los Estados Unidos y reducirían el precio de los productores estadunidenses, lo que los dejaría totalmente fuera del mercado del ácido sulfúrico. Las empresas canadienses firmaron contratos con varios productores estadunidenses que establecían que, a cambio de convertirse en distribuidores del ácido sulfúrico de las empresas canadienses, cada productor limitaría su propia producción de ácido sulfúrico y las empresas canadienses los compensarían por esto vendiéndoles ácido sulfúrico (para reventa) a un costo suficientemente bajo para volver más rentable la distribución que la producción. Estos acuerdos de "cierre" fueron impugnados como ilegales en sí mismos, debido a que un acuerdo para restringir la producción y,

[20] In re Sulfuric Acid Antitrust Litigation, 703 F.3d, 1004 (7th Cir., 2012).

por ende, aumentar el precio es el delito de antimonopolio ilegal en sí mismo de la fijación de precios.

La impugnación falló. Debido a que las empresas estadunidenses que estarían distribuyendo el producto de las empresas canadienses también eran productoras, éstas tenían una inquietud legítima con el efecto potencial sobre sus ventas de la producción de aquéllas. Si esa producción continuara, la oferta total de ácido sulfúrico en los Estados Unidos, aumentada como lo haría por la producción canadiense, superaría la demanda, provocando que los precios se desplomaran… quizás a un nivel en el que las empresas canadienses no pudieran recuperar el costo de la venta de su ácido en los Estados Unidos. Una inquietud relacionada era la posibilidad de la distribución de sus ácidos mediante empresas que también eran competidoras debido a que producían su propio ácido sulfúrico. No es ilegal en sí mismo insistir en que un distribuidor se comprometa a no vender una línea de productos que compita con la línea del proveedor; ¿qué diferencia debiera hacer que el mismo distribuidor produzca la línea de la competencia?

De este modo, el efecto de sostener que los acuerdos de cierre son ilegales en sí mismos podría haber sido el de disuadir a las empresas canadienses de entrar en el mercado de los Estados Unidos, en cuyo caso el precio del ácido sulfúrico en los Estados Unidos sería mayor debido a que la oferta total sería menor.

Sin embargo, los acuerdos de cierre crearon, de hecho, un cártel de productores de ácido sulfúrico, por lo que el caso plantea la cuestión de si debe haber una categoría de cárteles lícitos, porque son eficientes. Supongamos que las empresas competidoras formaran una agencia de distribución exclusiva, la cual decidirá el precio que se cobre por el producto de las empresas. Las empresas argumentan que la agencia de distribución exclusiva: *1)* reducirá los costos de búsqueda del comprador, *2)* aumentará los incentivos para innovar, y *3)* reducirá los costos esperados de la carga de peso muerto de la quiebra (véase XV.5 después). Éstos no son argumentos frívolos. La objeción para darles cualquier peso es dual: la extensa experiencia con la cartelización tanto en el extranjero como en muchas industrias en los Estados Unidos que siempre estuvieron (algunas todavía lo están) exentas de la ley antimonopolio ha revelado pocos o ningún ejemplo de cárteles que, en definitiva, promuevan la eficiencia; y sería administrativamente inviable para los tribunales llevar a cabo una investigación caso por caso sobre los costos y beneficios de los cárteles. Sin embargo, en el siguiente capítulo encontraremos que se sostiene que tal cártel "eficiente" (una agencia de distribución exclusiva formada por los competidores) es legal.[21] De modo que los cárteles son ilegales en sí mismos, susceptibles sólo a muy pocas excepciones *ad hoc*.

[21] Broadcast Music, Inc. *vs.* Columbia Broadcasting System Inc., 441 U.S. 1 (1979), analizado más adelante en XI.8.

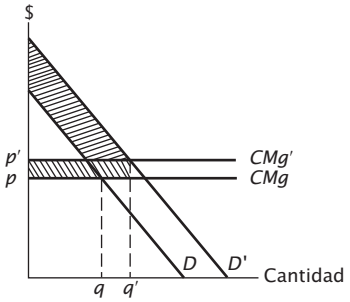
X.4. Mantenimiento del precio de reventa

Cuando la ley no lo prohíbe, los fabricantes fijarán a menudo un precio de reventa por debajo del cual no se permite que los detallistas u otros revendedores vendan su producto. En el caso de la Dra. Miles, la Corte Suprema sostuvo que, puesto que el resultado es el mismo que si los detallistas se reunieran y se pusieran de acuerdo sobre el precio a cobrar, lo que sería una fijación de precios ilegal, el mantenimiento de un precio de reventa es también ilegal en sí mismo.[22] Este razonamiento no explica la razón por la que el fabricante desearía que sus distribuidores detallistas tuvieran ganancias monopólicas, y omite la posibilidad de que el mantenimiento de un precio de reventa tenga un efecto adicional que no está presente en el caso del cártel de distribuidores: la provisión de un nivel óptimo de servicios de preventa para el producto del fabricante. Supongamos que haya un servicio de preventa que sea valioso para el consumidor y que sea provisto con mayor eficiencia por el minorista que por el fabricante, y la fijación de cuyo precio no pueda separarse de manera factible del precio del producto del fabricante: una sala de exposición bien provista, atendida por vendedores corteses y uniformados, sería un ejemplo.[23] El detallista que provee este servicio voluntariamente podría verse superado por un detallista competidor que no lo haga. El segundo detallista puede gorronear al primero, instando a sus consumidores a que visiten al primer detallista y regresen luego con él para obtener un precio de ganga posibilitado por el hecho de que él no hiciera el gasto de proveer los elaborados servicios de preventa provistos por el primer detallista. Al proveer servicios al consumidor sin cargo alguno, el primer detallista confiere un beneficio a los detallistas rivales que distribuyen la marca del fabricante.

Un fabricante puede impedir el gorroneo fijando un precio mínimo de menudeo a un nivel que genere ingresos, por encima de los costos mínimos de la distribución, suficientes para que el detallista provea el nivel de servicios que el fabricante desea. Impedido para competir por el precio, pero libre para competir en el servicio, cada detallista invertirá en servicios de preventa diseñados para arrebatarles consumidores a sus rivales. Esta competencia conti-

[22] Dr. Miles Medical Co. *vs.* John D. Park & Sons Co., 220 U.S., 373 (1911). Finalmente, el caso de la Dra. Miles se anuló en Leegin Creative Products, Inc. *vs.* PSKS, Inc., 551 U.S., 877 (2007). Véase también evidencia de que, al contrario del caso de la Dra. Miles, mantener el precio de reventa rara vez se emplea como método de colusión, en Pauline M. Ippolito, "Resale Price Maintenance: Empirical Evidence from Litigation", 34 *J. Law and Econ.*, 263 (1991).

[23] Pocas veces sería factible cobrarle a un cliente una tarifa para entrar en una sala de exposición, ya que para éste sería muy difícil determinar el valor del servicio que está recibiendo. Y, una vez que se crea la sala de exposición, se capacita a los vendedores, etc., el costo de informarle a otro cliente sobre lo que está a la venta es prácticamente nulo. Sin embargo, estos argumentos no son aplicables para los servicios de posventa, como la reparación.



GRÁFICA X.1

nuará hasta que los servicios de preventa hayan llegado al nivel deseado por el fabricante y se hayan eliminado por completo las ganancias monopólicas que los detallistas habrían ganado si hubiesen vendido al precio fijado sin ofrecer ningún servicio. Éste es otro ejemplo de la tendencia, discutida en el capítulo anterior, de las ganancias monopólicas potenciales a transformarse en costos (aquí los costos de los servicios de preventa) a través de la competencia en variables distintas del precio.[24]

No es que sea fácil *probar* que el bienestar económico aumenta por el mantenimiento del precio de reventa. Esto puede verse en la gráfica X.1, donde $CMg = p$ es el costo mínimo de la distribución detallista y es por ende el precio de menudeo cuando no se proveen servicios de preventa; p' es el costo mínimo de la distribución detallista fijado por el fabricante; CMg' es el nuevo costo de la distribución, que incluye el servicio de preventa que p' induce a los detallistas a ofrecer, y D' es la nueva y más alta curva de demanda creada por la provisión de servicios.

[24] La teoría de los "servicios de distribuidor" del mantenimiento del precio de reventa provoca las siguientes interrogantes adicionales:

a) ¿Presupone la teoría que el fabricante tiene poder monopólico? No, pero sí da por sentado que las marcas son lo suficientemente distintas para que los servicios proporcionados por el detallista de una marca no tengan el mismo valor para todas las demás.

b) ¿Cómo podríamos distinguir empíricamente entre el mantenimiento del precio de reventa impuesto por y para el beneficio del fabricante y el mantenimiento del precio de reventa que oculta a un cártel instigado por los detallistas? La prueba de fuego sería si la producción disminuyó o aumentó como consecuencia del esquema de mantenimiento del precio de reventa. Si es el primer caso, sugiere la cartelización; si es el último, sugiere la prestación de servicios que han aumentado la demanda del producto. No es suficiente sólo preguntar si el producto suele venderse con servicios de preventa elaborados: los servicios podrían ser el *resultado* de los esfuerzos de los miembros del cártel de un distribuidor para acaparar, tanto como sea posible, a través de una competencia (de servicio) distinta a la del precio, las ganancias esperadas derivadas de la formación del cártel.

c) Finalmente, ¿hay alguna diferencia analítica entre el mantenimiento de un precio de reventa, ya sea la asignación de territorios exclusivos para los detallistas o la prohibición de que vendan a otros detallistas o a través de ellos? No.

El fabricante se encuentra en mejor situación porque está vendiendo más, y al mismo precio. (El precio que le cobra al detallista no se ve afectado, o por lo menos no se ve necesariamente afectado, por la fijación de un precio de reventa que engendra los cambios mostrados en la gráfica.) Algunos consumidores se encuentran en mejor situación —incluidos, por definición, todos los que compran el producto por primera vez en virtud de los servicios que se proveen con él—, pero otros consumidores se encuentran en peor situación: quienes pagan un precio mayor que antes de que se proveyeran los servicios de preventa, pero no obtienen el beneficio correspondiente porque valúan los servicios en una cantidad menor al aumento de precio que están pagando por el producto.

Los beneficios y los costos se comparan en las áreas sombreadas de la gráfica x.1. El área sombreada que se encuentra entre las curvas de demanda es el beneficio derivado del mantenimiento de un precio de reventa; el área sombreada que se encuentra entre las curvas del costo marginal es el costo de tal beneficio. Aunque el beneficio se muestra como superior al costo, la gráfica podría retrazarse fácilmente para hacer que el costo supere al beneficio. El beneficio que confiere un fabricante a los consumidores vendiendo a un precio menor al que podrían verse obligados a pagar bajo un sistema de discriminación perfecta del precio —en otras palabras, el excedente del consumidor— no influye en las decisiones del fabricante, porque, por definición, él no recibe tal beneficio. Es un beneficio externo, de modo que el fabricante podría tomar medidas que tengan el efecto incidental de reducirlo, si tales medidas aumentan sus ganancias.

Esto no debería plantear problemas de antimonopolio. Cada vez que un fabricante altera la calidad de su producto, desplaza la curva de demanda, y los clientes intramarginales podrían perder más de lo que ganan los consumidores marginales. Sin embargo, su enfoque tiene que estar en los consumidores marginales a menos que pueda capturar la totalidad o la mayor parte del excedente del consumidor derivado de la alteración. No sería factible utilizar la ley antimonopolio para tratar de proteger a los consumidores intramarginales, que, de cualquier modo, son un blanco para los vendedores rivales.

A veces se describe que el efecto causado por el mantenimiento de un precio de reventa de aumentar el precio pero también la calidad (la cual aumenta con los servicios de preventa ofrecidos gratuitamente con un producto) involucra un intercambio de ventajas y desventajas entre la competencia de precios y de variables distintas del precio, lo que es correcto; o entre la competencia dentro de una misma marca y entre marcas distintas, lo que no es correcto. El mantenimiento de un precio de reventa limita en verdad la competencia de precios entre los detallistas de la marca cuyo precio de reventa ha fijado el fabricante —competencia de precios dentro de la marca—, y al hacerlo así promueve en efecto la competencia entre marcas al volver la marca del fabricante más atractiva en relación con otras marcas. Pero la compe-

tencia dentro de la marca no disminuye; simplemente cambia de la competencia de precios dentro de la marca a la competencia de servicios dentro de la marca. El mantenimiento de un precio de reventa no funcionará para un fabricante si no induce a sus detallistas a competir vigorosamente entre sí… pero a hacerlo en la provisión de servicios y no en la reducción de los precios.

Un fabricante podría imponer el mantenimiento de un precio de reventa a fin de asegurarse la lealtad y honestidad de sus detallistas garantizándoles una ganancia mayor de la que les daría la competencia pura, porque entonces la terminación de su contrato de distribución será una sanción más severa contra la deslealtad o la deshonestidad.[25] Esta justificación del mantenimiento de un precio de reventa no guarda ninguna relación con la teoría de los servicios del distribuidor; en efecto, es antitética porque supone que el distribuidor se embolsa el aumento de la diferencia entre el precio que paga a su proveedor y el precio que cobra a sus clientes, en lugar de que la competencia lo obligue a utilizar la diferencia para financiar los servicios adicionales del punto de venta; es por el embolso de la diferencia que es más fuerte la amenaza de liquidarlo en caso de que sea desleal: tiene más que perder si se le liquida.

X.5. Fusiones que crean monopolios u oligopolios

Sabemos que una ley que sólo castigaba la cartelización podía evadirse mediante la consolidación de los miembros del cártel en una sola empresa. Ésta es la justificación económica para la negativa de una mayoría de la Corte Suprema, en el caso de Northern Securities,[26] a adoptar la posición del magistrado Holmes en el sentido de que la Ley Sherman no era aplicable a las fusiones. Pero Holmes tenía razón para preocuparse por las implicaciones del alcance de la ley más allá de los cárteles para llegar hasta las fusiones. Las condiciones de la oferta y la demanda de un mercado podrían ser tales que una sola empresa pudiera abastecer toda la producción demandada a un costo medio menor que el de dos o más empresas; o una empresa podría tener una administración superior en cuyas manos los activos de todas las demás empresas valdrían más que ahora. Cualquiera de esas dos situaciones podría conducir a un monopolio a través de la fusión que acaso generara un ahorro de costos mayor a los costos sociales resultantes del precio monopólico, en cuyo caso la prohibición de la fusión sólo tendría un efecto a corto plazo. Puesto que, si las economías de escala decretan que no debe haber más de una empresa en el mercado, la competencia eliminará a todas salvo a una.

Desafortunadamente, resulta muy difícil distinguir fusiones que generan ahorros de costos de las fusiones que disminuyen la competencia. Sin em-

[25] Véase Benjamin Klein y Kevin M. Murphy, "Vertical Restraints as Contract Enforcement Mechanisms", 31 *J. Law & Econ.,* 265 (1988). *Cf.* después XII.7 y XII.8.

[26] Northern Securities Co. *vs.* United States, 193 U.S., 197 (1904).

bargo, una posibilidad consiste en examinar los movimientos de los precios de las acciones de los competidores de la empresa que se fusiona. Si los precios de sus acciones aumentan después del anuncio de la fusión, ello implicará que la fusión es conspirativa; si bajan, ello implicará que la fusión reduce los costos.

A pesar de su interés analítico, los monopolios son raros. Los oligopolios —mercados donde pocas empresas realizan la mayor parte de las ventas— son muy comunes, y la cuestión de su importancia competitiva es controversial. El número de empresas en un mercado es relevante para una preocupación por los cárteles, porque entre menor sea el número de empresas serán menores los costos de la coordinación de sus políticas sin un acuerdo explícito. El uso profiláctico de la ley antifusiones para evitar la creación de condiciones favorables para dicha colusión se convierte en una política atractiva, sobre todo si no se puede atacar la colusión tácita de manera directa.

X.6. El poder de mercado o monopólico

En la evaluación de las fusiones entre competidores y también de las recusaciones de la conducta de los monopolistas, los tribunales se preguntan si la empresa resultante de la fusión, o el monopolista, tiene "poder monopólico" (o "poder de mercado", un término más sugestivo de una variable continua). Es probable que una respuesta afirmativa condene a la fusión o proporcione un predicado esencial para condenar el comportamiento del monopolista como excluyente. Por lo que es importante considerar la manera en que la economía puede dar significado al concepto de poder de mercado.[27]

Las empresas generan la producción en la que coinciden el ingreso marginal y el costo marginal. Supongamos, por simplicidad, que el costo marginal es constante. Si el vendedor es una empresa competitiva, en el sentido estricto de que afronta una curva de demanda horizontal, lo que significa que puede vender todo lo que produzca al precio del mercado sin afectar a ese precio, entonces $P = IMg$; cada unidad adicional vendida aumenta los ingresos de la empresa en el precio de la unidad. Pero si la curva de demanda tiene una pendiente negativa, que es un prerrequisito para que la empresa tenga poder para afectar el precio con la variación de su producción, la venta de una unidad adicional de producción rendirá menos que P en ingreso adicional, porque la venta, al aumentar la producción, hará que P baje. ¿Cuánto? Esto depende de la elasticidad del precio de la demanda (e), lo que, recordemos, es el efecto sobre el precio de un pequeño incremento porcentual en la cantidad demandada, o el efecto sobre la cantidad demandada de un pequeño

[27] Lo que sigue amplifica el análisis breve de la economía básica de la competencia en I.1 anterior.

cambio porcentual en el precio. Si, por ejemplo, la elasticidad-precio de la demanda es 3 (en efecto –3, pero podemos olvidarnos del signo negativo), un aumento de 1% en la producción conducirá a una disminución de 3% en el precio. El ingreso marginal es igual a P $(1 - 1/e)$.[28] Dado que $IMg = CMg$, puesto que el precio competitivo sería igual a *CMg*, y ya que suponemos que *CMg* es constante, podemos escribir (con un poco de reordenamiento) la razón del precio monopólico al competitivo, P/CMg, como $e/(e - 1)$. En palabras, entre mayor sea la elasticidad de la demanda al precio monopólico maximizador de la ganancia, menor será la razón del precio monopólico al competitivo, y menor poder monopólico (o de mercado) tendrá la empresa. (La razón por la que e debe ser mayor a 1 en el caso del precio monopólico es que, por debajo de 1, un pequeño aumento porcentual en el precio provocará una reducción porcentual mucho menor en la cantidad demandada, por lo que el ingreso aumentará aunque los costos estarán cayendo porque se produce menos. Por ello, un aumento en el precio beneficiará al vendedor, y éste continuará elevándolo hasta encontrarse en la región elástica (es decir, $e > 1$ de la curva de demanda). Si e es 3, el precio monopólico será 50% mayor que el precio competitivo;[29] si es 2, será el doble del precio competitivo.

Esta fórmula es útil para demostrar que el poder monopólico: *1)* es variable antes que constante, y *2)* depende de la elasticidad de la demanda afrontada por la empresa en su precio maximizador de la ganancia. Otra fórmula nos permite derivar esa elasticidad (llamémosla e_{df}) de la elasticidad de la demanda afrontada por el mercado del que forma parte la empresa (e_{dm}), la elasticidad de la oferta de las otras empresas del mercado (e_s) y la participación de la empresa en el mercado (S). Esta fórmula es $e_{df} = e_{dm}/S + e_s(1 - S)/S$. Si una empresa tiene 100% del mercado, la elasticidad de la demanda afrontada por la empresa es, por supuesto, lo mismo que la elasticidad de la demanda afrontada por el mercado. Entre menor sea la participación de la empresa en el mercado, mayor será la elasticidad de la demanda que afronte en relación con la elasticidad de la demanda afrontada por el mercado.

Éste es un resultado cuya base lógica es de entrada patente. Es posible que no haya sustitutos cercanos para los adminículos, en cuyo caso su precio podría tener que aumentar mucho antes de que hubiera una sustitución sustancial de otros productos; por lo tanto, supongamos que e_{dm} es sólo 2, de

[28] Por lo que toca a la derivación de esta fórmula, véase William M. Landes y Richard A. Posner, "Market Power in Antitrust Cases", 94 *Harv. L. Rev.,* 937, 983-986 (1981). Para ayudar a entender la intuición detrás de la fórmula, nótese que si e fuera infinita, lo que significa que la empresa enfrenta una curva de demanda perfectamente horizontal (de modo que si acaso la empresa redujera su precio en total su producción sería infinita —una suposición obviamente poco realista—, mientras que si lo aumentara toda su producción caería a cero —algo más realista—), P sería igual a *CMg* y la empresa no tendría ningún poder monopolístico.

[29] La media del costo excesivo del cártel se ha estimado en 25%. John M. Connor y Robert H. Lande, "How High Do Cartels Raise Prices? Implications for Optimal Cartel Fines", 80 *Tul. L. Rev.*, 513, 559 (2005).

modo que si todos los productores de adminículos redujeran su producción 2% el precio aumentaría sólo 1%. Pero un productor que sólo venda 10% de los adminículos producidos por el mercado no podrá provocar un aumento de 1% en el precio reduciendo su propia producción 2%; tendrá que reducir la producción entera en el mercado 2%, lo que significa que debe reducir su propia producción 20%. Por lo tanto, e_{df} sería 20 aunque los competidores de la empresa no respondieran a su reducción de la producción aumentando la suya propia. Pero, a menos que estuvieran coludidos con esa empresa, ellos aumentarían su producción porque el precio de mercado había aumentado.

En cuánto aumentarían su producción dependería de la elasticidad de la oferta (la sensibilidad de la cantidad ofrecida ante un pequeño aumento en el precio). Supongamos que fuera de 1, lo que significa que un aumento de 1% en el precio los haría aumentar su producción 1%. Entre mayor sea su participación relativa en el mercado ($(1 - S)/S$), mayor será el efecto de su producción adicional sobre el precio. Con el supuesto de que es = 1, y de que la empresa que aumentó su precio tiene sólo 10% del mercado, es fácil mostrar que, tras tomar en cuenta la respuesta de la oferta, e_{df} no es 20, sino 29.

X.7. Definición del mercado y participaciones en el mercado

La sección anterior reveló cuáles son los determinantes del poder monopólico o el poder de mercado: la elasticidad de la demanda del mercado, la participación en el mercado de la empresa o las empresas que se suponen dotadas de poder monopólico, y la elasticidad de la oferta de otras empresas. Pero queda la tarea de estimar las elasticidades de la demanda y la oferta. Aunque los economistas han logrado algunos progresos en la estimación directa de estas elasticidades,[30] la mayoría de los tribunales en los casos antimonopólicos continúan recurriendo a aproximaciones muy burdas, resumidas en el concepto del mercado de producto y geográfico. Si un producto no tiene sustitutos cercanos en la demanda (es decir, nada que parezca proveer al consumidor los mismos servicios aproximadamente al mismo precio), y los vendedores de otros productos no pueden virar rápidamente para realizar su producción, entonces se supone que son bajas las elasticidades de la demanda y la oferta del mercado. De aquí puede inferirse que cualquier empresa con una partici-

[30] Véanse, por ejemplo, H. S. Houthakker y Lester D. Taylor, *Consumer Demand in the United States: Analyses and Projections,* 166-167 (2ª ed., 1970) (tabla 4.2); Robert Archibald y Robert Gillingham, "A Decomposition of the Price and Income Elasticities of the Consumer Demand for Gasoline", 47 S. *Econ. J.*, 1021 (1981); Eric A. Hanushek y John M. Quigley, "What Is the Price Elasticity of Housing Demand?", 62 Rev. Econ. & Stat., 449 (1980); C. Richard Shumway y Hongil Lim, "Functional Form and U.S. Agricultural Production Elasticities", 18 *J. Agric. & Resource Econ.*, 266 (1993); Hillard G. Huntington, "Inferred Demand and Supply Elasticities from a Comparison of World Oil Models", en *International Energy Economics*, 239 (Thomas Sterner [comp.], 1992).

pación sustancial en las ventas de este producto tiene un poder de mercado no trivial… a menos que los vendedores del producto idéntico en otras áreas geográficas pudieran entrar al mercado (local, regional o nacional) si el precio aumentara ligeramente allí, en cuyo caso el mercado se ha definido demasiado estrechamente y las elasticidades relevantes se han subestimado.

En el caso que cuestiona la monopolización del celofán de parte de Du Pont, la Corte Suprema sostuvo que el del celofán no era un mercado relevante porque había una alta elasticidad cruzada de la demanda entre el celofán y otros materiales de empaque flexibles.[31] Aun así, como sabemos, un monopolista vende siempre en la región elástica de su curva de demanda (porque la demanda inelástica implica que una elevación del precio no provocará una caída proporcional en la producción); y una razón por la que las curvas de demanda tienen una región elástica es que, entre mayor sea el precio de un producto, más atractivos se vuelven los productos sustitutos para el consumidor. Es posible que la elevada elasticidad cruzada de la demanda de celofán haya significado sólo que la Du Pont no podría haber aumentado más el precio del celofán sin perder gran cantidad de ventas a manos de productos sustitutos, y esto sería consistente con el hecho de que el precio existente fuese el precio monopólico. El énfasis por parte de la Corte en una elevada elasticidad cruzada de la demanda habría tenido más sentido si el caso hubiera implicado una consolidación de los productores de celofán, donde el problema habría consistido en saber si la consolidación crearía o aumentaría el poder monopólico. Si la elasticidad cruzada de la demanda entre el celofán y los materiales de empaque sustitutos era elevada antes de la consolidación, la empresa consolidada tendría escaso poder para aumentar el precio.[32]

Pero esto conduce a la paradoja de que una fusión de duopolistas podría no ser ilegal porque, mediante la colusión tácita, habían aumentado *ya* el precio del mercado hasta el punto en el que los sustitutos pobres habían pasado a formar parte de su mercado; pero esto pasaría por alto el hecho de que es probable que la fusión aumente el nivel de precios en el mercado, porque es probable que la colusión por parte de los duopolistas sea imperfecta (cada uno tendrá un incentivo para engañar al otro reduciendo el precio a los clientes grandes) en comparación con el monopolio.

Debido a los altos costos de transportación en relación con el valor de un producto, no todos los fabricantes del producto podrían competir por los

[31] United States *vs.* E. I. Du Pont de Nemours & Co., 351 U.S., 377 (1956).

[32] No se sigue de aquí lo contrario: si, en nuestro caso hipotético de fusión del celofán, fuese baja la elasticidad cruzada de la demanda entre el celofán y otros materiales de empaque, no tendríamos derecho a concluir que la demanda de celofán fuese inelástica al precio. El hecho de que un producto no tenga sustitutos cercanos no implica necesariamente que los consumidores estén dispuestos a pagar precios mayores antes que prescindir de ese producto. Todo lo que se necesita para asegurar que la demanda de un producto sea elástica al precio es que los consumidores tengan usos alternativos para su dinero, a los que recurrirán en gran número si aumenta el precio relativo del producto por encima del nivel competitivo.

mismos clientes; en otras palabras, los mercados pueden estar geográficamente limitados. Los tribunales han tendido a incluir en el mercado a los vendedores que venden efectivamente al mismo grupo de clientes y excluir a quienes no lo hacen. Esto es a la vez demasiado y muy poco. Si el mercado está monopolizado, el precio monopólico atraerá a vendedores de mercados distantes que no podrían haber cubierto sus costos de transportación y otros costos de venta si se hubiese cobrado el precio competitivo (el problema del celofán). Si el mercado no está monopolizado, podría haber un grupo de vendedores distantes que por el momento no enviaran su producto al mercado pero que podrían hacerlo si el precio aumentara incluso ligeramente (es decir, hay buenos sustitutos en la producción). Es posible que los costos de los vendedores externos en la venta a este mercado fuesen 2% mayores que los de los vendedores internos a causa del gasto de transportación. Esto implicaría que, si a resultas de la monopolización aumentara 2% el precio del mercado, los vendedores externos empezarían a enviar su producto al mercado y el precio no podría aumentar más. Es posible que incluso los altos costos de transportación no otorguen un poder monopólico a los vendedores cercanos. Supongamos que los vendedores lejanos incurren en costos de transportación adicionales iguales a 6% de sus costos totales, pero que sus costos distintos de la transportación se encuentran 4% por debajo de los costos de los vendedores cercanos; entonces los obstáculos de transportación que enfrentan los vendedores lejanos no serían mayores que en nuestro ejemplo anterior.

Si los vendedores distantes realizan algunas ventas locales al precio competitivo, ¿cómo se calcularán las participaciones en el mercado? Supongamos que los vendedores locales fabrican y venden 90 de las 100 toneladas de adminículos vendidos en Illinois; una empresa de Indiana vende las otras 10. Sin embargo, la empresa de Indiana vende en otra parte el doble de los adminículos que vende en Illinois; su producción total es entonces de 30. Suponiendo que el precio actual del mercado de Illinois es el precio competitivo,[33] ¿cuál es la participación de la empresa de Indiana en el mercado de Illinois: 10% (10/100) o 25% (30/30 + 90))? La respuesta es 25%. Dado que la empresa de Indiana ha superado todos los obstáculos que enfrenta por las barreras del costo de transportación derivadas de su distancia de los consumidores de Illinois, podría desviar el resto de su producción hacia Illinois si el precio aumentara incluso ligeramente… y lo haría. Pues debe de estar ganando tanto en Illinois como en otras partes; de otro modo desviaría sus ventas de Illinois hacia otra parte ahora mismo. Un aumento del precio en Illinois haría que

[33] Ésa es una calificación importante. Si el precio de mercado actual está por encima del nivel competitivo, puede afectar e involucrar a los vendedores del mercado que no estarían allí si el precio de mercado fuera competitivo debido a que sus costos de suministro del mercado serían más altos que los de las empresas "dentro" del mercado. Así, la presencia de los precios de monopolio podría crear la ilusión de que el mercado es más grande de lo que realmente es y más pequeña la participación de mercado del demandado.

sus ventas allí fueran más rentables, de modo que tendría un incentivo para desviar sus ventas hacia ese mercado. Esta limitación del poder monopólico de las empresas de Illinois se capta incluyendo en el mercado de Illinois las ventas totales del vendedor de Indiana, no sólo las ventas que realiza en ese estado. Adviértase la implicación de que dos productos están en el mismo mercado si sus precios están correlacionados estrechamente de manera positiva.[34] La correlación sugiere que cualquier diferencia de precio se eliminará inmediatamente por sustitución hacia o lejos del producto cuyo precio ha cambiado.

La participación de mercado en un mercado bien definido puede ser un indicador de poder monopólico, pero esto no significa mucho. Sabemos, por ejemplo, que si la elasticidad de la demanda en el mercado es 2 y la elasticidad de la oferta es 0, una empresa con 50% del mercado afrontaría una elasticidad de la demanda de 4, y esto le permitiría cobrar un precio 33% por arriba del precio competitivo, lo que es mucho. Pero, si su participación fuese de sólo 20%, todavía podría cobrar un precio 11% mayor que el precio competitivo. No obstante, dado que una mayor elasticidad de demanda en el mercado, o una alta elasticidad de la oferta, podrían reducir grandemente esta cifra, resulta difícil basar cualesquiera conclusiones sólo en la proporción en el mercado. Y esto omite la probabilidad sustancial de que si una empresa ha adquirido un gran tamaño independientemente de las fusiones recientes probablemente sea más eficiente que sus competidores. Pues si la empresa se creó o amplió por las fusiones en el pasado remoto, entonces tendría que mantener su precio bajo en relación con los precios cobrados por los competidores reales y potenciales para conservar su participación de mercado. Sólo podría hacer eso si fuera más eficiente que los otros. De ser así, sus costos más bajos podrían ser mayores que los costos sociales derivados del cobro de un precio monopólico.

El potencial del análisis económico para hacer posibles las estimaciones más precisas del poder monopólico se ilustra por el enfoque utilizado por la Comisión Federal de Comercio (FTC) en su esfuerzo exitoso para impedir que la cadena de tiendas de artículos de oficina Staples adquiriera a Office Depot.[35] Staples y Office Depot son dos de las tres cadenas de supertiendas de artículos de oficina (la tercera es OfficeMax). Comoquiera que sea, las tres tenían, aun considerándolas en conjunto, sólo una participación moderada en las ventas totales de menudeo de artículos de oficina. La Comisión sostuvo que, sin embargo, las tres constituían un mercado separado en virtud de la mayor conveniencia que ofrecían a los compradores. Si se hubiese permitido que se

[34] Véase George J. Stigler y Robert A. Sherwin, "The Extent of the Market", 28 *J. Law & Econ.,* 555 (1985).

[35] FTC *vs.* Staples, Inc., 970 F. Supp., 1066 (D.D.C., 1997). Véanse Serdar Dalkir y Frederick R. Warren-Boulton, "Prices, Market Definition, and the Effects of Merger: Staples-Office Depot", en *The Antitrust Revolution: Economics, Competition, and Policy*, 143, 153 (John E. Kwoka Jr. y Lawrence J. White [comps.], 3ª ed., 1999).

realizara la fusión, se habría reducido el número de supertiendas de tres a dos en algunas localidades, y de dos a una en otras (aquellas donde no estaba presente OfficeMax).

La Comisión presentó pruebas de que los precios cobrados por las supertiendas eran mucho menores que los precios cobrados por otros vendedores detallistas de artículos de oficina. Con todo, el hecho de que algún producto se venda a precios diferentes en tipos de tiendas diferentes no prueba que haya dos mercados; es posible que la tienda de precio mayor ofrezca más servicios, y la diferencia de precios podría reflejar simplemente el costo adicional de tales servicios, del mismo modo que el precio de un producto de marca, mayor que el de su marca privada equivalente, puede compensar simplemente un costo de comercialización mayor. Pero las supertiendas estaban ofreciendo *mayor* conveniencia (en particular una variedad mucho mayor) y un precio menor. Esto probaba que otros detallistas no eran sustitutos cercanos para las supertiendas.

El perito de la Comisión estudió las regresiones de los precios cobrados por las supertiendas con respecto a diversas variables explicativas posibles, incluyendo la presencia de una o más de las otras supertiendas en el mercado local particular. La regresión reveló que los precios cobrados por Staples y Office Depot eran mucho menores en las localidades donde estaban presentes las tres supertiendas, y algo menores en las localidades donde estaban presentes dos de ellas, que en aquellas donde sólo estaba presente Staples u Office Depot. Esto implicaba que los precios cobrados por cada una de las empresas en fusión estaban constreñidos por otras supertiendas a un nivel significativamente menor de lo que lo estaban por otros detallistas de artículos de oficina en los mercados donde sólo había una supertienda. Podría parecer extraño que esos mercados no atrajeran la creación de nuevas supertiendas, pero muchos de los mercados no tenían demanda suficiente para soportar supertiendas adicionales, y esto volvería aleatorias a las entradas nuevas: el nuevo miembro tendría que derrotar a una supertienda existente para poder permanecer en el mercado. La Comisión realizó también un estudio de eventos que demostró que la evaluación por parte del mercado de valores de la fusión propuesta se basaba en una expectativa de que la fusión permitiría en efecto que la empresa resultante de ella cobrara precios mayores.

X.8. La competencia potencial

La competencia potencial es un concepto importante del derecho antimonopólico… o, mejor dicho, dos conceptos: la competencia potencial *percibida* y la competencia potencial *efectiva*. El primero capta mejor la idea de la competencia "potencial". Las empresas que no venden en un mercado pero que lo harían si el precio de mercado fuese mayor son competidores potenciales. Las empresas que *entrarán* al mercado en el futuro aunque el precio no au-

mente —los "competidores potenciales efectivos" en la jerga antimonopólica— se describen mejor simplemente como los competidores futuros.

Sabemos que a mayor elasticidad de la demanda que afronta una empresa, menor poder de mercado que posee, y que si un aumento del precio provocará nueva producción de otras empresas —es decir, si la elasticidad de la oferta es positiva—, la elasticidad de la demanda de la empresa será mayor que en otras circunstancias. Sin embargo, esto sugiere que no hay necesidad de una doctrina separada de la competencia "potencial". Todo lo que se requiere es una definición de los mercados suficientemente amplia para incluir a las empresas que, aunque no venden ahora en el mercado en cuestión, lo harían si el precio aumentara ligeramente. Supongamos que el alambre de aluminio y el de cobre son buenos sustitutos en la producción porque se usan las mismas máquinas en la producción de ambos, pero que son malos sustitutos en el consumo. Si el precio del cobre aumentara por encima del nivel competitivo, los productores del alambre de aluminio se desplazarían hacia la producción de alambre de cobre; y su capacidad para el desplazamiento se aproxima por su producción actual de alambre de aluminio: cuanto más alambre de aluminio produzcan, podrían producir más alambre de cobre sin una expansión costosa de las instalaciones de producción. Por lo tanto, la producción del mercado de alambre de aluminio debiera incluirse en el cálculo de la participación de mercado de las partes de una fusión entre un productor de alambre de cobre y uno de aluminio.

Pero éste es un caso en el que el nuevo miembro no tiene que construir nuevas instalaciones. ¿Qué ocurriría si tuviera que hacerlo? Dado que se requiere tiempo para la construcción de instalaciones de producción, la amenaza de entrada planteada por tal empresa afectará más probablemente a la elasticidad de la demanda afrontada por las empresas del mercado a largo plazo que a corto plazo; y, dado que la colusión es en gran medida un fenómeno de corto plazo porque un precio conspirativo provocará tanto el engaño como la atracción de empresas nuevas al mercado, es posible que la eliminación de tales amenazas no sea suficientemente importante para ameritar la preocupación antimonopólica, sobre todo porque será difícil calcular las participaciones de mercado de las empresas que no tienen todavía ninguna capacidad productiva. En efecto, será muy difícil identificar a las empresas que tienen probabilidades de construir una capacidad productiva para entrar al mercado si el precio de mercado aumenta por encima del nivel competitivo.

Además, la aplicación de la doctrina de la "competencia potencial percibida" a tales empresas supone que las empresas que se coluden elegirían un precio que impidiera la entrada incluso a largo plazo. Esto se llama limitación del precio, y no está claro en modo alguno que tenga sentido económico.[36]

[36] Pero véase Margaret A. Peteraf y Randal Reed, "Pricing and Performance in Monopoly Airline Markets", 37 *J. Law & Econ.,* 193 (1994).

(Si no lo tiene, tampoco lo tendrá el intento de preservar competidores potenciales a fin de promover la limitación del precio.) Dado que es probable que los costos a largo plazo de un nuevo miembro no sean mayores a los costos de las empresas que ya se encuentran en el mercado, la limitación del precio implica imponer un precio igual o cercano al costo marginal: una política cuestionable desde el punto de vista de la maximización de las ganancias, sobre todo porque es probable que la colusión se derrumbe por su propio peso en breve tiempo. Es probable que la estrategia de maximización de las ganancias de quienes se coluden sea una de fijación del precio sobre la base de la elasticidad de la demanda (menor) a corto plazo, olvidando así a todos los competidores potenciales que no pueden entrar al mercado de inmediato porque carecen de las instalaciones necesarias.

Esto deja aún en pie la posibilidad de que haya competidores futuros, cuya entrada podría reducir el precio a largo plazo, aunque la percepción de su entrada potencial no tuviera ningún impacto a corto plazo. Pero, si es poco probable que la colusión persista más allá del corto plazo, ¿por qué preocuparse por la eliminación de un competidor futuro mediante la fusión? ¿Como un profiláctico contra los cárteles futuros? Pero, dado que las ganancias competitivas futuras valen menos que las ganancias presentes a causa del descuento hasta el valor presente, ¿cuán probable es que los costos de la cartelización futura superen a cualesquiera ahorros de costos (los que se obtendrían antes) posibilitados por la fusión?

X.9. La depredación

Hasta ahora hemos hecho hincapié en la adquisición de poder monopólico mediante la consolidación u otras formas de la cooperación entre los competidores. Una cuestión importante consiste en saber si tal poder podrá obtenerse o aumentarse mediante los esfuerzos de una empresa. Podemos omitir el caso en el que la empresa obtiene un monopolio o una participación mayor en el mercado gracias a su eficiencia superior o a una licencia gubernamental, concentrando nuestra atención en las tácticas que se consideran abusos del proceso competitivo. Una de ellas es la discriminación depredadora de los precios: una empresa vende por debajo del costo en algunos mercados y, después de que sus competidores han sido expulsados, fija un precio monopólico.[37] Menos dramáticamente, la empresa vende por debajo del costo para disciplinar a alguien que baja su precio o para evitar nuevas entradas al mercado.

[37] Podríamos preguntarnos: ¿por qué si los consumidores se verán afectados a largo plazo, cooperarán con el depredador comprándole cuando venda por debajo del costo? La razón es que hay un problema de externalidad de grandes cantidades: cada consumidor sabe o percibe que abstenerse de negociar con el depredador tendrá un efecto mínimo sobre el éxito del esquema depredador.

Los casos confirmados de discriminación depredadora del precio eran raros aun antes de que la práctica fuese ilegal. La práctica es muy costosa para el depredador, quien incurre en una pérdida presente y sustancial en aras de ganancias que no sólo son diferidas sino que podrían ser temporales porque, una vez que los competidores existentes sean expulsados del mercado y se establezca un precio monopólico, nuevos competidores se verán atraídos al mercado por ese precio; la táctica del precio depredador podría tener que repetirse. Además, entre más grande sea el monopolista por comparación con el nuevo miembro, más cara será la venta por debajo del costo para el monopolista que para el neófito, porque el monopolista tendrá que reducir el precio de toda su producción, a menos que pueda discriminar por precio, otorgando descuentos sólo a los clientes objetivo del recién ingresado.

Si la mera *amenaza* de la discriminación depredadora de precios fuese suficiente para mantener a raya a los competidores, se emplearía a menudo porque cuesta menos la amenaza (excluyendo los costos del castigo legal). Pero una amenaza debe ser creíble para que sea eficaz. Una amenaza de vender por debajo del costo no lo sería de ordinario. La víctima de la amenaza sabría que el amenazador estaría restringido por su propio interés para cumplirla, porque es muy costoso vender por debajo del costo. Sin embargo, la amenaza puede ser creíble cuando el amenazador tiene una posición monopólica en varios mercados diferentes mientras que cada uno de sus competidores vende en uno solo de tales mercados. El monopolista podría ser capaz de convencer a cada competidor de que cumplirá una amenaza de vender por debajo del costo en mercados particulares a fin de volver más creíbles sus amenazas en otros mercados. Incluso podría ser capaz de impedir nuevas entradas fijando un precio menor a su costo la primera vez que una empresa entre a su mercado; esto podría ser suficiente para desalentar entradas subsecuentes a cualquiera de sus mercados. El costo de uno o dos episodios de venta por debajo del costo podría ser pequeño por comparación con los beneficios que obtiene el depredador de haber establecido su credibilidad.[38]

Contrario a esta conclusión, cabe señalar que un nuevo miembro tiene menos que perder cuando entra a una escala limitada, y que es más probable que el monopolista tolere un desafío limitado para su posición que uno general. El neófito limitado evitará también todas las deseconomías causadas por la operación a una escala tan grande como la que el monopolista está obligado a emplear a fin de mantener su monopolio, por lo que el neófito podría tener costos más bajos, lo que aumentará el costo para el monopolista de reducir el precio a un nivel suficientemente bajo para aplastarlo. Tampoco está claro que el hecho de vender por debajo del costo unas cuantas veces volvería creíble la amenaza del monopolista de continuar haciéndolo si vuelve

[38] Véase una analogía en el análisis de los gastos litigiosos presentado después en XXII.11.

a ser retado. Es posible que sólo esté fanfarroneando; en efecto, nuestra discusión sugiere que *está* fanfarroneando, que está vendiendo por debajo del costo una o dos veces porque cree que de ese modo no tendrá que seguir haciéndolo, no porque tenga un compromiso irrenunciable para enfrentarse a cada nuevo miembro o competidor engañoso con precios por debajo del costo.

Si el monopolista pudiera crear de algún modo una reputación de sentir una ira irracional, implacable contra quienquiera que tuviera la osadía de entrar a un mercado en el cual él fuera el único vendedor, o de vender a un precio menor al suyo (la clase de reputación que crean los hombres de honor en una sociedad en la que la venganza es la base del control social; véase antes VIII.5), su amenaza de ejercer represalias contra cualquier neófito vendiendo por debajo del costo por mucho dinero que perdiera en el proceso ya no sería un mero alarde y no tendría que cumplir su amenaza. Pero precisamente porque una reputación de estar dispuesto a ejercer represalias contra una infracción de nuestros derechos percibidos, cualquiera que sea el saldo de costos y beneficios, es una buena inversión, resulta difícil que alguien que desee tal reputación convenza a todo el mundo de que no es realmente una persona racional que simula la irracionalidad. Sabemos de I.4 que, dado un conjunto finito de nuevos miembros potenciales, la amenaza del monopolista de vender por debajo del costo no será creíble a menos que los neófitos potenciales crean que el monopolista es racional. Sin embargo, si hay incertidumbre acerca de su racionalidad, su amenaza podría ser creíble aunque en efecto él sea racional. Más aún: dada esa incertidumbre, el monopolista puede cumplir la amenaza aunque fuera racional, a fin de crear una reputación de irracionalidad.[39] Por lo tanto, el establecimiento de precios depredadores puede ser una estrategia racional si al menos algunos vendedores grandes son irracionales pero no se sabe cuáles son.[40]

No obstante, otro procedimiento para volver creíble una amenaza, como sabemos del capítulo VIII, es el del compromiso previo.[41] Una estrategia de

[39] Véase Douglas G. Baird, Robert H. Gertner y Randal C. Picker, *Game Theory and the Law,* 181-183 (1994).

[40] Solamente un vendedor grande podría practicar los precios depredadores. Un vendedor pequeño no podría expandir su producción lo suficiente para causar mucho daño al nuevo miembro (es decir, forzando sustancialmente a la baja el precio de mercado), y es probable que tampoco fuera capaz de recuperar sus pérdidas, aumentando el precio por encima del nivel competitivo, si se expulsa al neófito; pues será poco probable que el vendedor más pequeño tenga poder de monopolio, y si la mayoría de los vendedores son pequeños la colusión será difícil.

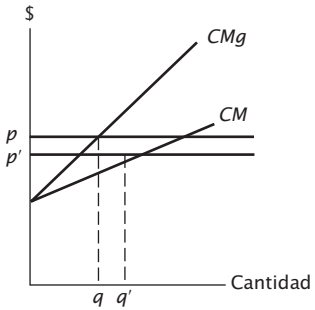
[41] Véase una discusión clásica del papel del compromiso previo en la teoría de juegos en Thomas C. Schelling, "An Essay on Bargaining", 46 *Am. Econ. Rev.,* 281 (1956). Véanse pruebas empíricas de que las empresas presuntamente depredadoras *no* adoptan una estrategia de compromiso previo, como los modelos de precios depredadores de la teoría de juegos implican que lo harían, en John R. Lott Jr. y Tim C. Opler, "Testing Whether Predatory Commitments Are Credible", 69 *J. Bus.,* 339 (1996).

compromiso previo plausible para un monopolista consiste en construir más capacidad que la necesaria para satisfacer su demanda actual o prevista para el futuro. Entonces, si entra al mercado una empresa nueva, el monopolista, al tener una capacidad excedente, podrá expandir su producción a un costo muy bajo, lo que le permite vender a menor precio que el nuevo miembro sin ir por debajo de sus propios costos marginales.[42] Sabiendo que el monopolista podría enfrentar la nueva entrada al mercado en esta forma relativamente indolora (y probablemente legal), los neófitos potenciales podrían ser disuadidos. Por supuesto, la estrategia le costaría algo al monopolista, ya que estaría pagando por una capacidad que quizá no usaría nunca.

Si una empresa puede ahuyentar una nueva entrada vendiendo igual o por encima de sus propios costos, esto implica que es la empresa más eficiente, y no querríamos llamarla depredadora. Supongamos que el depredador potencial posee poder monopólico, de modo que está vendiendo por encima de su costo marginal y no a este costo, y decide bajar su precio. Mientras que no baje su precio más allá del costo marginal de su nuevo y más alto nivel de producción, generado por la reducción del precio, no estará vendiendo por debajo del costo en ningún sentido relevante para la política de la competencia, porque su reducción del precio no podría sacar del mercado a nadie que no fuese un competidor menos eficiente.

El problema de convertir a esta observación en la base de una regla legal es que el costo marginal no es una cifra que se lleve en los libros contables de una empresa o que pueda derivarse fácilmente de las cifras que sí se encuentran allí (lo que señalamos antes, al discutir los daños contractuales en IV.10). Es cierto que los costos marginales son una función de los costos variables antes que de los costos fijos —por definición, los costos fijos no se ven afectados por los cambios del nivel de producción—, y por lo general los costos variables se determinan fácilmente. Sin embargo, los costos marginales y los costos variables no son sinónimos. Supongamos que el costo de la mano de obra y los materiales y otros costos variables de la producción de 100 adminículos ascienden a 100 dólares, y que serían 99 si sólo se produjeran 99 unidades, de modo que el costo marginal de una empresa a un nivel de producción de 100 es 1 dólar. Pero supongamos además que si se aumenta la producción a 101 unidades, presionando la capacidad existente de la empresa, sus costos variables totales de toda su producción se dispararán a 110 dólares. Por lo tanto, su costo variable medio será 1.09 dólares (110 dólares ÷ 101). Pero su costo marginal será 10 dólares. Si esta empresa bajara su precio de 3 a 2 dólares, a fin de crear una demanda para la mayor producción, estaría siguiendo una estrategia de precios depredadores, pero esto sería difícil de probar porque un tribunal tendría dificultades para calcular el costo marginal de la

[42] El costo marginal tenderá a ser bajo cuando la producción pueda expandirse sin construir nuevas plantas.



GRÁFICA X.2

empresa. Aunque se ha sugerido que se utilice el costo variable medio en tal caso,[43] este costo es en nuestro ejemplo, como hemos visto, de sólo 1.09 dólares. Su empleo como una aproximación al costo marginal (que es 10 dólares) conduciría a la conclusión incorrecta de que la rebaja del precio (a 2 dólares) no era depredadora, ya que 2.00 > 1.09 dólares.

La gráfica X.2 es una representación esquemática de este problema. La empresa aquí representada baja el precio a p' y aumenta su producción a q' (el par p' y q' representa el nivel precio-producción que obstaculiza o destruye a las empresas rivales al menor costo para el depredador). En este punto, su costo marginal supera a su precio. Pero su costo variable medio, que aumenta más lentamente que el costo marginal (debido a que el alto costo de la última unidad de producción se promedia con los costos menores de las unidades intramarginales), se encuentra por debajo de su precio, creando la impresión engañosa de que la empresa no está siendo depredadora.

Otro problema es que el cálculo del costo variable y, por ende, el del costo marginal es muy sensible al periodo de tiempo. A muy corto plazo, la mayor parte de los costos son fijos, mientras que a muy largo plazo virtualmente todos son variables (una excepción sería el costo de la creación de una sociedad anónima). La renta, por ejemplo, es un costo fijo a corto plazo pero variable a largo plazo, pues si cambia la necesidad de espacio de la empresa pagará una renta diferente tras la expiración del arrendamiento actual. Lo mismo se aplica a los seguros, los sueldos de los ejecutivos, los pagos de pensiones, los impuestos a la propiedad, la depreciación, la publicidad y muchos otros costos.

En principio, el periodo dentro del cual debe determinarse si los costos son fijos o variables debiera ser el periodo de la presunta práctica de precios depredadores. Si la reducción del precio dura un mes, los costos variables pertinentes serán los que varían en el curso de un mes, como la mano de obra

[43] Phillip Areeda y Donald F. Turner, "Predatory Pricing and Related Practices under Section 2 of the Sherman Act", 88 *Harv. L. Rev.,* 697 (1975).

pagada por hora que se emplea en la fabricación del producto. Entre más dure la disminución del precio, mayor será la parte de los costos de la empresa que son variables. Si se permitiera que una empresa fijara indefinidamente su precio al nivel de su costo variable a corto plazo, podría expulsar del mercado a una empresa más eficiente, simplemente porque los costos variables de esa empresa a largo plazo serían mucho mayores incluso que los costos variables a corto plazo de una empresa menos eficiente. De hecho, a largo plazo, que es el periodo pertinente cuando se persiste durante largo tiempo en una rebaja del precio presuntamente depredadora, los costos variables medios tienden a fusionarse con los costos medios totales (costos fijos más costos variables, todos divididos por la producción de la empresa). En el muy largo plazo, son muy pocos los costos fijos, es decir, inevitables: ni siquiera el alquiler, los seguros, los impuestos o las pensiones.

¿Pero tendría sentido alguna vez que una empresa fijara precios depredadores durante un periodo prolongado? Después de todo, ello alargaría el periodo de reposición —cuando el depredador, habiendo destruido a su presa, eleva su precio hasta un nivel monopólico—, y las ganancias monopólicas obtenidas durante ese periodo tendrían que ser descontadas a su valor presente al decidir si la venta a menos del costo (un costo corriente) será benéfica.

Podría parecer que un precio por debajo del costo marginal *a corto plazo* nunca pudiera ser depredador. El argumento sería que, a menos que la empresa tenga capacidad excedente, su costo marginal a largo plazo debe ser menor que su costo marginal a corto plazo, porque el productor tiene mayor flexibilidad para ajustar sus costos a los cambios en la demanda en el largo plazo que en el corto plazo; por ejemplo, en el largo plazo el productor puede remplazar su planta actual por una de mayor capacidad si aumenta la demanda, pero en el corto plazo está limitado por la configuración de su capacidad actual. Por lo tanto, un precio igual al costo marginal en el corto plazo no podría excluir a un competidor igualmente o más eficiente, como lo haría un precio por debajo del costo marginal en el largo plazo. Sin embargo, la calificación "a menos que" es importante. Recordemos de VI.7 anterior que la mayoría de las empresas tienen alguna capacidad de reserva, la cual es un "excedente" en cierto sentido, pero no así un derroche. Una empresa podría adquirir capacidad de reserva deliberadamente para satisfacer los picos inesperados de la demanda o inevitablemente a causa de lo abultado de los activos. En ambos casos, el costo marginal a corto plazo será menor que el costo marginal a largo plazo hasta que se agote esa capacidad excedente por la expansión de la producción. Porque, hasta que se llegue a ese punto, la empresa no tendrá que rentar un espacio adicional, contratar nuevo personal o tomar prestado más dinero — éstos serán costos fijos antes que marginales a corto plazo—, aunque podría incurrir en algunos costos adicionales porque tiene que pagar salarios extra o comprar materias primas en el mercado de presentes.

El potencial para fijar precios al nivel del costo marginal en el corto plazo a fin de expulsar a un competidor igual o más eficiente aumenta si el competidor opera a una escala mucho menor. Por ejemplo, si para el monopolista es normal la operación al 90% de la capacidad (siendo el otro 10% una capacidad de reserva antes que excedente), y el monopolista es 10 veces mayor que el competidor, el monopolista podrá aumentar su producción lo suficiente para abastecer a todos los clientes de sus competidores sin toparse con el límite de su capacidad, aunque sólo a corto plazo.

Supongamos que una empresa fabrica muchos productos diferentes, y algunos de los insumos —el tiempo de sus ejecutivos, por ejemplo— son los mismos para los diferentes productos. Si la empresa baja el precio de uno solo de sus productos, ¿cómo debieran tratarse los sueldos de los ejecutivos, a corto y a largo plazos, al decidir si la rebaja del precio es depredadora? Los contadores, preocupados como deben estarlo por la minimización de la discreción de los administradores en la caracterización de los gastos, prefieren a menudo reglas estrictas antes que los matices económicos. De modo que requieren que los gastos de publicidad se traten como gastos corrientes aunque, dado que los efectos de la publicidad duran a menudo más de un año, en realidad son gastos de capital y debieran depreciarse a lo largo de su vida útil. Esto podría hacer la diferencia en el caso de la fijación de precios depredadores si los gastos de publicidad se registran como un gasto, por un lado, o si se capitalizan y amortizan, por el otro. Si los gastos en publicidad aumentan, registrarlos como un gasto conducirá a una mayor deducción que si se capitalizan y amortizan, en cuyo caso la deducción será de una cifra promedio a lo largo de varios años y, por lo tanto, será menor.

Por último, un estándar legal basado en el costo marginal o en el costo variable medio es manipulable porque, dentro de ciertos límites, una empresa puede sustituir los costos variables por costos fijos, incluso sin construir un exceso de capacidad. Supongamos que la empresa puede elegir entre una planta muy intensiva en capital y una planta muy intensiva en mano de obra. La primera tendrá mayores costos fijos, la última tendrá mayores costos variables. Por lo tanto, la primera ruta dará a la empresa una mayor flexibilidad para la fijación del precio bajo un estándar legal que equipare la fijación de precios depredadores a la venta por debajo del costo marginal o del costo variable medio.

X.10. Ejecución hipotecaria, ataduras y agrupamientos, barreras a la entrada

Además de la fijación depredadora de los precios, varias otras prácticas que según se dice son creadoras de poder monopólico tienen también la propiedad de imponer al monopolista costos que serán a menudo por lo menos tan elevados como los que tienen que soportar las víctimas en mente. Suponga-

mos que un fabricante compra todas las tiendas de menudeo que venden el producto de su industria a fin de excluir a sus competidores del acceso al mercado. Éstos reaccionarán construyendo sus propias tiendas. Esto implicaría un costo para ellos, pero no mayor que el costo para el monopolista potencial de comprar una cadena de tiendas de menudeo que tendría un gran exceso de capacidad después de que sus rivales construyeran sus propias tiendas.

Tal adquisición, que reúne empresas en etapas de la distribución sucesivas (antes que competidoras), es una fusión "vertical". A menudo esto puede aliviar, más que exacerbar, los problemas de poder monopólico. Considérese una fusión entre dos monopolistas que ocupan posiciones sucesivas en la cadena de distribución y recuérdese que la distribución puede conceptualizarse como un costo para el productor medido por la diferencia existente entre el precio de mayoreo y el de menudeo, que esta diferencia será mayor si el distribuidor es un monopolista y que el precio cobrado por un monopolista es mayor entre mayor sea el costo marginal del monopolista. Por lo tanto, en el caso de monopolios sucesivos, la fusión, al reducir el costo de distribución del productor, se traduce en un precio menor (aunque todavía monopólico) y en consecuencia en una producción mayor.

Considérese ahora la fusión del productor de un producto final con el productor monopólico de un insumo que (al revés de la distribución) se emplea en proporciones variables con otros insumos para producir el producto final. Si el productor monopólico del insumo cobra un precio monopólico al productor del producto final, este último sustituirá ese insumo empleando más de otro insumo —la distorsión típica provocada por el monopolio—. Pero si el monopolista del insumo adquiere al productor del producto final, se elimina esta posibilidad de sustitución y se genera una combinación de insumos más eficiente.

A menudo se ha considerado a la ejecución hipotecaria como un método eficaz para la obtención de poder monopólico por parte de una empresa que tiene ya un monopolio en un mercado relacionado. Supongamos que una empresa tiene una patente para un producto, digamos computadoras, y que se niega a vender o rentar su producto a menos que el comprador o arrendatario acepte adquirir de ella sus abastos de algún otro producto, digamos discos duros. La empresa obtiene así un monopolio de los discos (el producto atado) empleados con sus computadoras (el producto atador). Pero no obtiene ganancias monopólicas del segundo monopolio. Si cobra a sus arrendatarios de computadoras un precio mayor que el precio competitivo por los discos, ellos tratarán esto como un aumento indirecto de la renta de la computadora (ya que están interesados en el costo total del servicio de cómputo que están obteniendo); pero, si están dispuestos a pagar una renta mayor, ello significa que la compañía de computadoras podría haber explotado directamente esa disposición, cobrando una renta mayor.

Sin embargo, una ventaja posible para un monopolista de imponer una atadura es que podría facilitarle la discriminación de precios. Si la compañía de computadoras cobra sus ingresos monopólicos mediante el precio que fija por cada disco, en efecto la tarifa de la renta de la computadora variará de un cliente a otro según la intensidad del uso de cada cliente, lo que podría ser una aproximación razonable a la elasticidad de la demanda de éste, siendo los usuarios mayores quienes tienen la elasticidad menor.[44] Pero desde este punto de vista la atadura no es un método por el que una empresa pueda emplear un monopolio en un mercado para obtener ganancias monopólicas también en un segundo mercado, ni es probable que cierre el acceso al mercado a los fabricantes rivales del producto atado. Dado que el monopolista del producto que ata no puede obtener ganancias monopólicas de la venta del producto atado, no le interesará controlar su fabricación. Sólo le interesa que sus ventas del producto atado se canalicen a través de él. Así pues, los acuerdos de atadura no perturban necesariamente la estructura existente del mercado del producto atado. Sin embargo, todavía se podrían objetar, porque en la medida en que aumenten las ganancias esperadas del monopolio aumentan los recursos empleados en obtener y mantener los monopolios.

Una versión más sutil de la teoría de la ejecución hipotecaria sostiene que la ejecución hipotecaria aumenta el poder monopólico indirectamente al crear una barrera a la entrada. En sentido estricto, una barrera a la entrada es una condición que vuelve a los costos a largo plazo de un nuevo miembro a un mercado mayores que los costos a largo plazo de las empresas existentes; un ejemplo es una limitación reguladora de la entrada. Sin embargo, este término se emplea también como un sinónimo de grandes costos de iniciación de operaciones. De este modo, un mercado en el que son grandes los gastos de capital o de publicidad necesarios para obtener una posición firme se describe con frecuencia como aquel en el que hay barreras para la entrada, y es en este último sentido del término que la ejecución hipotecaria podría crear tal barrera.

Si las empresas existentes en el mercado fueran propietarias de todas las tiendas de menudeo y estuvieran decididas a negárselas a nuevos miembros, un neófito tendría que abrir sus propias tiendas de menudeo, lo que aumentaría sus requerimientos de capital. O, si los discos duros estuvieran atados a las rentas de computadoras, una persona que acabara de entrar en el negocio de

[44] Pero no una aproximación infalible, porque un comprador más grande puede tener más recursos para buscar un proveedor de menor costo y, por lo tanto, con una elasticidad de la demanda menor.

Adviértase que la atadura no se puede utilizar como un método de discriminación del precio si la atadura y los productos atados se emplean en proporciones fijas antes que variables (por ejemplo, un zapato derecho y uno izquierdo) para producir el producto final (un par de zapatos), porque en tal caso la atadura no sirve para monitorear la intensidad de uso que haga el cliente del producto atador. Todo el mundo compra los zapatos por par.

las computadoras tendría que arreglárselas también para vender los discos a sus clientes. Pero los costos de capital de las empresas existentes serían mayores también a resultas de la integración vertical, de modo que no está claro cuál ventaja obtendrían estas empresas. El neófito podría tener que pagar una tasa de interés mayor si su empresa se considerara riesgosa, pero no si se tratase de una empresa establecida garantizando su crédito total al prestamista. Y lo mismo ocurre con la publicidad. Un mercado en el que las empresas gastan mucho en publicidad podría ser de difícil acceso sin grandes gastos en publicidad, pero el neófito puede consolarse con el hecho de que las empresas del mercado deben incurrir también en grandes costos de publicidad a fin de mantener sus posiciones.[45] La publicidad es un mal ejemplo de una barrera a la entrada por otras tres razones. El nuevo miembro se aprovecha de la publicidad de las empresas originales del mercado, la que ha creado una aceptación pública del producto; y tiene la opción de gastar menos en publicidad y vender a un precio menor que el de las empresas existentes, confiando en que las grandes cadenas de detallistas publiciten la disponibilidad de un nuevo sustituto de precio bajo; y la publicidad vuelve más elástica la demanda en tanto que se orienta hacia los consumidores marginales. Pues tal publicidad involucra en el mercado a consumidores que no están comprometidos con el producto del anunciante (eso es lo que significa para ellos ser marginales), y así pueden ser inducidos fácilmente a cambiar a un competidor. Pero la publicidad puede crear lealtad a la marca del anunciante, especialmente si se orienta a consumidores inframarginales en lugar de simplemente a los marginales, como podría suceder: ya sea porque los dos objetivos no se pueden separar de inmediato o porque el anunciante quiere afianzar la preferencia de los consumidores inframarginales por su marca, frente a la de los nuevos miembros.

Si no preguntamos cuál sería la cuantía de los costos de iniciación de quien ingresa, sino cuál sería el monto de la razón existente entre los costos de iniciación y los costos de operación —en otras palabras, la razón de los costos fijos a los variables—, identificaremos un problema de entrada real. Una razón elevada de los costos fijos a los variables implica que el mercado podría tener espacio sólo para unas cuantas empresas —quizá sólo una empresa— de tamaño eficiente, lo que vuelve muy riesgosa la entrada porque, si el nuevo miembro vende menos de lo que espera, sus ingresos caerán pero sus costos no mucho, ya que son fijos. Éste es el problema del monopolio natural que se discute en el capítulo XIII.

Otra versión de la teoría de la ejecución hipotecaria sustituye el concepto de una barrera a la entrada por el de un factor que demora la entrada. Como mencionamos antes, los costos de entrada son, dentro de cierto intervalo, una función negativa del tiempo que se requiere para entrar (véase antes IX.6). Y el

[45] Por supuesto, esto depende de la rapidez de la depreciación de la publicidad.

periodo óptimo para efectuar la entrada es mayor entre más complejas sean las operaciones que deben coordinarse para iniciar la producción. La integración vertical podría aumentar la complejidad de la entrada al hacer que el neófito tenga que entrar a dos niveles en lugar de a uno solo (la distribución de menudeo y la fabricación en el ejemplo anterior). Pero, si una empresa se integra verticalmente a fin de desalentar la entrada, es probable que incurra en costos mayores (porque, si la integración vertical se justificara por el costo, es probable que la empresa ya se hubiera integrado verticalmente, es decir, sin la necesidad de un motivo depredador para hacerlo), lo que alentará la entrada.

Un vendedor podría ser capaz de obtener un monopolio simplemente persuadiendo a muchos clientes potenciales de firmar un contrato para negociar exclusivamente con él, hasta que los clientes restantes fuesen demasiado pocos como para financiar a otro vendedor que operara en una escala eficiente.[46] Cada cliente firmaría por una pequeña cantidad, sabiendo que si no lo hiciera podría acabar todavía afrontando a un monopolista, ya que la cooperación de todos los clientes no es esencial para el plan del monopolista potencial. (Compárese con las ofertas de compra de dos niveles que se discuten en el capítulo xv.) En la ejecución hipotecaria por contrato antes que por la propiedad, el monopolista potencial no incurre en los costos, y las posibles deseconomías de escala o de alcance, de la expansión a una industria diferente; debe pagar algo a sus clientes para inducirlos a renunciar a sus alternativas, pero posiblemente muy poco si cada uno de los clientes sabe que su aquiescencia no es esencial.

Un ejemplo famoso, aunque quizá mal entendido, de ejecución hipotecaria por contrato es el arreglo celebrado entre la Standard Oil Trust de John D. Rockefeller y los ferrocarriles que transportaban petróleo en cuyos términos los ferrocarriles otorgaban rebajas a la Standard Oil, lo que tenía el efecto de elevar los costos de transportación de otras compañías petroleras, impidiéndoles así que compitieran en términos de igualdad con el monopolio. Se ha sostenido que el monopolio era en efecto el agente del cártel de los ferrocarriles, que a cambio de ayudar a los ferrocarriles a mantener precios por encima del nivel competitivo al abstenerse de obtener descuentos secretos (como podría haberlo hecho el monopolio) que habrían propagado la desconfianza entre los cartelistas, el monopolio era recompensado con rebajas de parte de los ferrocarriles que le otorgaban una ventaja competitiva sobre otras compañías petroleras.[47]

El tema de la ejecución hipotecaria se pone de relieve en un juicio donde se acusó de monopolización ilegal al productor monopólico de maquinaria

[46] Eric B. Rasmusen, J. Mark Ramseyer y John S. Wiley, Jr., "Naked Exclusion", 81 *Am. Econ. Rev.*, 1137 (1991).

[47] Elizabeth Granitz y Benjamin Klein, "Monopolization by 'Raising Rivals' Costs': The Standard Oil Case", 39 *J. Law & Econ.*, 1 (1996).

para fabricar zapatos.[48] La United Shoe Machinery Corporation se negaba a vender sus máquinas a sus clientes (fabricantes de zapatos); sólo las rentaría, y por un plazo mínimo de 10 años. El tribunal de distrito falló que esta política perjudicaba las posibilidades de los productores rivales impidiéndoles entrar, durante los periodos de los contratos, al gran mercado integrado por los clientes de la United. No se considera la posibilidad de que un fabricante rival de maquinaria de zapatos pudiera reunir a clientes de la United antes de la fecha de expiración de los arrendamientos. Cada año, cerca de 10% del mercado estaría en libertad para contratar con el fabricante rival. Y, dado que la vida útil de la maquinaria de zapatos (tanto su vida física como, durante el periodo en cuestión en cualquier caso, su vida económica) es mayor de 10 años, un fabricante de zapatos que comprara las máquinas de la United, en lugar de rentarlas, estaría en todo caso raras veces en el mercado para una máquina nueva hasta que hubiesen transcurrido más de 10 años desde que compró su máquina actual. En tal caso, un contrato de arrendamiento por 10 años podría no ser más restrictivo que una venta incondicional.

Sería improbable que los clientes de la United se unieran a una campaña para fortalecer su posición monopólica sin insistir en ser compensados por la pérdida de fuentes de abasto alternativas y menos costosas (por ser competitivas). Las empresas reparten a menudo sus compras entre proveedores rivales a fin de preservar la competencia en sus mercados de abasto. Lo hacen aunque, como vimos en el caso de los precios depredadores, haya un problema de gorroneo. Supongamos que un fabricante de maquinaria de zapatos ofrece arrendar una máquina por 10 000 dólares al año bajo un contrato terminable a voluntad, mientras que la United ofrece arrendar una máquina similar por 9 000 dólares al año pero insiste en un arrendamiento de 10 años diseñado para destruir a los productores rivales y permitirle a la United elevar su precio a 20 000 dólares al término del periodo. El precio de 9 000 dólares no es una ganga para los fabricantes de zapatos. El trato ofrecido por la United impone un costo adicional al comprador, medido por el valor presente del precio mayor en el futuro. Si ese valor es, digamos, de 2 000 dólares al año, la oferta de la United equivale a cobrar 11 000 dólares al arrendatario. Cada fabricante de zapatos podría aceptar todavía la oferta de la United, por pensar ya sea que otros fabricantes de zapatos celebrarían contratos de arrendamiento con el segundo productor de maquinaria por 10 000 dólares, manteniendo así a esa empresa en el mercado y protegiéndolo contra un monopolio de la United, o bien que le arrendarían a la United por 9 000 dólares, dejándolo con costos mayores si pagara 10 000 dólares al segundo productor pero sin protección para el futuro, a menos que el segundo productor pudiera sobrevivir sólo con su patronazgo.

[48] United States *vs.* United Shoe Machinery Corp., 110 F. Supp., 295 (D. Mass.), *affirmed per curiam*, 347 U.S., 521 (1953).

Pero si la gente pensara siempre de ese modo ningún cártel sería siquiera parcialmente eficaz, a menos que fuese legalmente exigible. Cada miembro del cártel vendería por debajo del precio de la organización, calculando que todos los demás estaban haciendo lo mismo o que todos los demás estaban vendiendo al precio del cártel; en ambos supuestos, el engaño es rentable. Así como los cárteles no son viables de ordinario si es muy grande el número de empresas que tendrían que unirse para volverlo eficaz, es improbable que sea eficaz la resistencia de los clientes de un monopolista a la tentación del gorroneo si los clientes son muy numerosos. Pero la United no estaba vendiendo a centenares o millares de clientes, sino a un número pequeño de fabricantes de zapatos.

Además, el segundo productor en nuestro caso hipotético podría superar el problema del gorroneo ofreciendo arrendar la máquina por 10 000 dólares, a condición de que un número mínimo especificado de fabricantes de zapatos le arrendaran sus máquinas. Cada arrendatario consideraría ventajoso celebrar tal contrato en lugar de pagar 9 000 dólares y aceptar una condición de exclusividad de la United, ya que el costo anual del arrendamiento de la United es en realidad de 11 000 dólares.

Existe una motivación no excluyente para una política de sólo arrendar un bien muy durable tal como la maquinaria de zapatos. Como se explicó antes en IX.6, quien compra tal bien a un monopolista no estará seguro de que los compradores subsecuentes, que podrían ser sus competidores, no obtengan un precio menor; es probable que los compradores subsecuentes tengan una demanda del producto más elástica, ya que ésa es una razón probable de que ellos no hubiesen sido los primeros en comprar el bien en cuestión. El problema puede resolverse si el monopolista arrienda el bien en lugar de venderlo, siempre que la renta se ligue al valor del bien para el comprador. Porque entonces, si el monopolista disminuye ese valor aumentando su producción y vendiendo a un precio menor la producción adicional a los competidores del arrendatario, disminuirá su ingreso como arrendador.

Otra posibilidad no excluyente es que la United estuviera tratando de impedir el desarrollo de un mercado de segunda mano. Aun cuando la durabilidad de un producto, incluido el precio que de seguro alcanzará finalmente en el mercado de segunda mano del producto, es algo por lo que el comprador estará dispuesto a pagar, los compradores de ese mercado que terminaran con el bien podrían no evaluarlo tan alto como lo hacen muchos compradores corrientes. En ese caso, el esfuerzo del monopolista para cobrar un precio elevado a los compradores corrientes se verá frustrado por la disposición de los dueños de segunda mano a prescindir del producto a un precio menor.[49]

[49] Véase Michael Waldman, "Eliminating the Market for Secondhand Goods: An Alternative Explanation for Leasing", 40 *J. Law & Econ.,* 61 (1997). Pero adviértase que esto podría considerarse como excluyente desde el punto de vista de los distribuidores de segunda mano.

Al arrendar en lugar de vender, el monopolista puede impedir el desarrollo de un mercado de segunda mano.

Ambos métodos de explotación de un monopolio de un bien durable son, como la atadura, métodos para el aumento de las ganancias monopólicas, pero no para excluir a los competidores.

Otra práctica excluyente alegada en el caso de la United Shoe Machinery fue la del cobro de un solo precio por la maquinaria y su reparación. El tribunal consideró que este "agrupamiento" había retardado el surgimiento de una industria independiente de reparación de maquinaria de zapatos y así volvió más difícil que otros fabricantes de esta maquinaria compitieran con la United, ya que ésta les estaba cobrando a sus clientes por el servicio de reparación independientemente de que lo utilizaran o no, lo que desalentaba al cliente de buscar en otra parte el servicio de reparación para sus máquinas de la United. Pero el hecho de obligar a otros productores de maquinaria de zapatos a contratar servicios de reparación sólo aumentaría sus costos en relación con los de United en el caso en que existieran economías de escala en la reparación.

El agrupamiento es analíticamente similar a la atadura. En el agrupamiento se regala el producto "atado", lo que significa que en efecto el vendedor está pagando al comprador para que lo compre. En la atadura se obliga al comprador a comprar el producto atado como una condición para obtener el producto atador. Se parece al dilema de la zanahoria contra el garrote, pero en realidad se trata de dos zanahorias, ya que un vendedor no puede obligar efectivamente a la gente a hacer nada. El agrupamiento es el equivalente del ofrecimiento de un descuento del precio fijado al producto atador. Pero la atadura es lo mismo. A fin de inducir al comprador a comprar algo extra que no se desea (si se desea, no hay necesidad de una atadura) o algo extra que se desea a un precio por encima del nivel competitivo (el caso en el que se emplea la atadura como un método de discriminación de los precios), el vendedor debe ofrecer en el producto atador un trato mejor que el que ofrecería de otro modo; podría incluso regalar el producto base.

La discriminación de precios es el propósito habitual del agrupamiento y de la atadura. Supongamos que una empresa vende dos productos, X e Y, que tienen valores diferentes para consumidores diferentes. Supongamos que la empresa tiene dos clientes para estos productos, A y B, y que A pagaría 8 000 dólares por X y 2 500 por Y, mientras que B pagaría 7 000 dólares por X y 3 000 por Y. Si la empresa fijara por separado el precio de X y el de Y, su mejor precio para X sería 7 000 dólares y para Y 2 500, de modo que su ingreso total sería 19 000. Pero, si vende X e Y en paquete, podrá cobrar 10 000 dólares por el paquete y obtener así un ingreso total de 20 000 dólares.[50] La razón

[50] Véase George J. Stigler, "A Note on Block Booking", en Stigler, *The Organization of Industry*, 165 (1968).

es que cuando se fijan los precios de los productos por separado el precio es disminuido por el comprador que valúa cada bien menos que el otro comprador; el agrupamiento elimina este efecto. La rentabilidad del agrupamiento es mayor entre mayor sea el número de los productos que puedan agruparse.[51] Porque esto vuelve más probable que el paquete contenga productos a los que los consumidores asignan valuaciones opuestas, como en el ejemplo numérico, donde A valúa X más que B, mientras que B valúa Y más que A.

El agrupamiento y la atadura se realizan generalmente por contrato, pero no necesariamente. El agrupamiento puede asumir literalmente la forma de una fusión de componentes, de modo que el consumidor no podría usar un componente obtenido de otro proveedor, y la atadura puede asumir la forma de una configuración del producto atador en tal forma que sólo el producto atado provisto por el vendedor del producto base funcionaría bien con él, dejando al consumidor sin otra opción que comprarle al vendedor el producto atado.

En principio, el agrupamiento y la atadura pueden excluir a un nuevo miembro igualmente eficiente, o más, obligando al neófito a entrar con un producto más complejo que en ausencia de tales estrategias. Siempre que tal entrada sea más riesgosa y, por ende, más costosa (en lugar de que requiriera simplemente más capital al mismo costo unitario del capital), la entrada se verá frenada o impedida. En el caso del agrupamiento, el cliente que desee el resto de lo agrupado se resistirá a pagar dos veces por uno de los productos agrupados comprándolo a otro vendedor también.[52]

Así pues, la atadura de piezas de reparación o repuesto a la compra del equipo original podría ser concebiblemente una práctica excluyente; pero esto es improbable, y la Corte Suprema tuvo dudosamente razón cuando pensó que el de Eastman Kodak Co. *vs*. Image Technical Services, Inc. era tal caso.[53] A lo sumo, Eastman Kodak estaba explotando a sus consumidores aprovechándose de los costos que ellos habían hundido en la compra original de su fotocopiadora de alto volumen y su equipo micrográfico para obligarlos a comprarle partes de repuesto a un precio exorbitante. Es cierto que la Kodak podría hacer esto sólo si fuera un "monopolio" en cierto sentido, aunque un monopolio limitado a los compradores de su equipo original. (No tenía poder monopólico en el mercado de equipo original.) Pero, excepto bajo la competencia perfecta, una abstracción de libro de texto aunque aproximada en ciertos mercados agrícolas, todos los vendedores tienen compradores "cauti-

[51] Véase Yannis Bakos y Eric Brynjolfsson, "Bundling and Competition on the Internet", 19 *Marketing Sci.,* 63 (2000).

[52] Barry Nalebuff, "Bundling as an Entry Barrier", 119 *Q. J. Econ.,* 159 (2004); Nalebuff, "Exclusionary Bundling", 50 *Antitrust Bull.,* 321, 365 (2005).

[53] 504 U.S., 451 (1992). Véanse Benjamin Klein, "Market Power in Antitrust: Economic Analysis after *Kodak*", 3 *Supreme Ct. Econ. Rev.,* 43 (1993); Herbert Hovenkamp, "Market Power in Aftermarkets: Antitrust Policy and the *Kodak* Case", 40 *UCLA L. Rev.,* 1447 (1993).

vos" a corto plazo. El vendedor que explota su "monopolio" de piezas de repuesto se encontrará sin muchos compradores de su equipo original en el periodo siguiente.

Otra explicación para la política de precios de la Kodak en cuanto a las piezas de repuesto, y más generalmente en cuanto a la percepción generalizada de que las piezas de repuesto provistas por el fabricante original del equipo tienen un "precio excesivo", es que la Kodak y otros fabricantes de productos de marca están practicando la discriminación de precios. Los productos que venden dichos fabricantes son productos diferenciados, cada uno de ellos con una curva de demanda de pendiente negativa. A los compradores que tienen una demanda relativamente inelástica a causa de los altos costos de tiempo se les puede cobrar un precio mayor por el producto al cargárseles un precio elevado por las piezas de repuesto y por el servicio, porque ellos, al revés de los compradores cuyo tiempo no es valioso, prefieren la conveniencia de volver al fabricante del equipo original por las piezas de reposición y el servicio, antes que buscar las alternativas disponibles en talleres de reparación más baratos pero menos confiables.

Otra práctica cuestionada como excluyente es el "descuento de fidelidad", ilustrado por LePage's Inc. *vs*. 3M (Minnesota Mining and Manufacturing Co).[54] 3M fabricó la cinta Scotch, que hasta principios de la década de 1990 tenía cerca de 90% del mercado para la cinta transparente. En la década de 1980 LePage's había iniciado la fabricación de cinta transparente para la venta bajo la marca privada de los minoristas. Su cinta hizo avances considerables en las ventas de cinta Scotch y 3M respondió comenzando la comercialización de cinta etiquetada con su propia marca, ofreciendo descuentos de fin de año a los minoristas que compraban varias líneas de productos de 3M. Esta oferta dio a los minoristas un fuerte incentivo para comprar la cinta transparente de la marca privada 3M en lugar de la de LePage's. La cuota de mercado de LePage's se redujo y entonces demandó. Y perdió; ¿debió perder?

Los descuentos nunca presionaron el precio de 3M por debajo de su costo, ya fuese para la cinta Scotch o para la cinta de marca privada. Tampoco se obligó a ningún minorista a comprar la cinta de la marca 3M como condición para que se le permitiera comprar cinta Scotch. Pero, debido a que los minoristas consideraban indispensable para su negocio llevar la cinta Scotch, tenían un fuerte incentivo para comprar su cinta de marca privada así como la de 3M, ya que ello reducía el precio real que estaban pagando por la cinta Scotch. Mientras que los costos de la marca privada de LePage's fueran los mismos que los de 3M, LePage's podría haber igualado los descuentos de 3M sin tener pérdidas. Pero difícilmente habría sido posible para LePage's igualar los descuentos de 3M sobre la base de cliente por cliente. Habría tenido que hacer un corte uniforme de precios, lo que habría otorgado ganancias inespe-

[54] 324 F.3d, 141 (3d Cir. 2003) (en banc).

radas a esas tiendas que carecían de una fuerte demanda de cinta Scotch y que por lo tanto no recibían grandes descuentos de 3M, mientras que probablemente no habrían sido lo bastante extensos para igualar las rebajas recibidas por aquellas tiendas que sí tenían una fuerte demanda de cinta Scotch y que, de esta manera, obtenían grandes rebajas de 3M.

La corte consideró que 3M era un fabricante más eficiente que LePage's. Pero supongamos que fuera más eficiente no porque tuviera costos unitarios menores sino simplemente porque existen economías de escala en la producción de cinta transparente. Existe una diferencia entre la eficiencia basada en costos más bajos de la mano de obra o de los materiales, una administración superior, mejor calidad y otros atributos específicos de una empresa, por un lado, y, por otro, la eficiencia basada en la escala, la cual puede alcanzar cualquier empresa capaz de incrementar su producción a una escala eficiente. En la medida en que los descuentos de fidelidad elevaron los costos promedio de LePage's al reducir su producción e impedirle de ese modo alcanzar las economías de escala disponibles, esto no fue una consecuencia de que 3M fuese una empresa más eficiente, en un sentido relevante para la política antimonopolio. 3M pudo haber temido que, si se permitía a LePage's ponerse al tanto, minoristas poderosos como Office Depot y Wal-Mart preferirían vender una cinta más barata de marca privada que impulsar la cinta Scotch.

Una explicación alternativa, no excluyente, para los descuentos de fidelidad de 3M es que, al igual que en muchos casos de agrupamiento, eran una forma de discriminación de precios. Algunos minoristas valoraban más la cinta Scotch que la cinta de marca privada y otros al contrario. Los descuentos proporcionan, en efecto, un precio único para un paquete que consta de dos productos (no conozco los otros productos que cubren las rebajas). Los minoristas que prefieren la cinta Scotch verían los descuentos como una forma de poder comprar cinta Scotch más barata; quienes prefieren la cinta de marca privada los verían como aquellos que hacen que la cinta de marca privada se pueda adquirir más barata.

Otra posibilidad es que los descuentos de lealtad tienen como propósito inducir… a la lealtad. Otro nombre para la lealtad podría ser el de bajos costos de negociación. Una vez que el minorista se acostumbra a satisfacer sus necesidades de cinta transparente comprándole a 3M, cualquier recién llegado tendrá que recompensar al minorista por el riesgo y los (otros) costos de cambiar de proveedor. Eso desalentará la nueva entrada de empresas que no sean más eficientes que los vendedores existentes, pero por buenas razones económicas.

Un último punto para considerar es que, aunque las prácticas de exclusión se asocian por lo general con cárteles o con empresas individuales que tienen poder monopólico, también pueden ser el resultado de la colusión tácita entre competidores.[55]

[55] Véase C. Scott Hemphill y Tim Wu, "Parallel Exclusion", 122 *Yale L. J.*, 1182 (2013).

X.11. El antimonopolio y la "nueva economía"

La preocupación por las prácticas excluyentes es particularmente aguda en lo tocante a las empresas que operan en la "nueva economía", la que incluye industrias tales como la manufactura de software de computadora, la provisión de servicios por empresas basadas en la Internet, y los servicios y equipos de comunicaciones que apoyan a las dos primeras industrias. Las industrias de la nueva economía difieren marcadamente de la mayoría de las industrias en las que se desarrolló la doctrina antimonopólica: industrias como la fabricación de cigarrillos, acero y maquinaria de calzado, caracterizadas por la producción en múltiples plantas y múltiples empresas (lo que indica que las economías de escala son limitadas al nivel de la planta y de la empresa), demanda estable, gran inversión de capital, tasas de innovación moderadas, y entrada y salida lentas e infrecuentes. Por el contrario, las industrias de la nueva economía tienden a caracterizarse por los costos medios decrecientes (para los productos, no para las empresas) en un amplio intervalo de la producción, requerimientos de capital moderados, tasas de innovación muy elevadas, entradas y salidas frecuentes y economías de escala en el consumo, cuya realización podría requerir la cooperación monopólica o entre empresas en la fijación de estándares. El primer punto, los costos medios decrecientes, es resultado del hecho de que, con la excepción parcial del equipo de comunicaciones, la producción principal de las industrias de la nueva economía es la propiedad intelectual, es decir, el código de computadora, antes que bienes físicos. La creación de propiedad intelectual es a menudo cara, pero una vez creada es de ordinario bajo el costo de fabricación de copias adicionales; tan pasmosamente bajo en el caso del software que es sólo una ligera exageración tratar el costo marginal como igual a cero.

Las economías de escala en la fabricación son familiares; hasta cierto punto, entre más larga sea la corrida de producción, menor será el costo total medio a medida que los costos fijos se distribuyen entre un nivel de producción cada vez mayor. Las economías de escala en el consumo, más comúnmente llamadas externalidades de red, describen la situación en la que, entre mayor sea la producción de la empresa, hasta cierto punto, más valiosa será esa producción para sus consumidores. El servicio telefónico carece de valor si sólo hay un suscriptor, que no tiene con quien hablar. Entre más suscriptores haya, más valioso es el servicio para cada uno, o por lo menos para muchos de ellos. (Ésa es la *externalidad* de la red. Cada suscriptor adicional confiere un beneficio a los demás suscriptores que podrían desear ponerse en contacto con él.) Los servicios interactivos, tales como las subastas por correo electrónico y en línea, son similares. Lo mismo ocurre cuando se comparten programas de computadora, como cuando dos o más académicos colaboran en la redacción de un artículo científico por medio de programas de procesamiento de pala-

bras y hojas de cálculo. Aparte del trabajo en red literal o del compartir, los programas de computadora tienden a ser más valiosos entre más personas los usen porque se facilitan la capacitación, el apoyo del personal de tecnología de la información, y la estandarización de equipos y procedimientos. La primera y la tercera razones explican el hecho de que se haya estandarizado el teclado de la máquina de escribir.

Las externalidades de red presuponen la uniformidad antes que una fuente común. El sistema telefónico internacional es una sola red, pero sus componentes son propiedad de un vasto número de empresas e individuos separados. Los componentes han sido estandarizados para asegurar la interoperabilidad, del mismo modo que se ha estandarizado el ancho de las vías de ferrocarril. Pero el hecho de que los activos esenciales de las empresas de la nueva economía sean formas de la propiedad intelectual, que pueden estar monopolizadas legalmente, y los servicios provistos por estas empresas se caractericen por externalidades de red, empuja a la nueva economía hacia el monopolio, pero también, curiosamente, hacia la competencia. Recuérdese que la competencia para obtener un monopolio es una forma importante de la competencia (el "huracán de destrucción creativa" de Schumpeter). Entre mayor sea la protección de que goce contra la competencia la empresa que logre obtener un monopolio, mayor competencia habrá para convertirse en tal monopolista; y siempre que los únicos medios viables o permitidos para la obtención del monopolio sean socialmente productivos, esta competencia podría ser enteramente deseable. Una empresa que tendrá la protección de la ley de propiedad intelectual y de las economías de escala en el consumo si es la primera en fabricar un componente esencial de un producto o servicio de la nueva economía tendrá un monopolio especialmente lucrativo, y esta perspectiva debiera acelerar la tasa de innovación, así como, en igualdad de las demás circunstancias, entre más valioso sea un tesoro enterrado, más rápidamente será recuperado.

Más aún, es probable que el monopolista de éxito de la nueva economía sea una empresa que cobra inicialmente un precio muy bajo por el producto nuevo que ha creado. Volvamos al ejemplo del teléfono. Dado que cada suscriptor nuevo aumenta el valor del servicio para los suscriptores existentes, una compañía telefónica tiene un incentivo para proveer precios atractivos a los suscriptores nuevos; el dinero que perderá con ellos puede ser recuperado con creces por el precio mayor que pagarán los suscriptores existentes por el acceso a una red más grande. Esto es especialmente probable si la red será un monopolio natural, en el sentido de que a ningún competidor le resultaría viable su duplicación; entonces, entre más rápidamente alcance la red la madurez, el monopolista estará protegido durante más largo tiempo contra los desafíos a su monopolio. Pero el desafío al monopolista de red existente, de ordinario mediante una innovación revolucionaria (una "aplicación asesina"), debe ser viable para que funcione el huracán de la destrucción creativa; y el monopo-

lio podría desalentar la innovación de otras empresas. Las externalidades de red podrían otorgar al monopolista una ventaja de costos que supere al beneficio de una tecnología nueva superior. Éste es un ejemplo de la "dependencia de la ruta": Es posible que una industria esté atorada con una tecnología inferior porque la ventaja de costos de la red existente no puede ser superada.[56]

Las prácticas que otorgan a un monopolista de red mayor aislamiento frente a la competencia de lo que es inherente a su control de la red preocupan particularmente a la política antimonopólica. Consideremos como analogía el caso antiguo de Standard Fashion.[57] El demandado fabricaba una popular línea de patrones de diseño que las mujeres podían usar para hacer sus propios vestidos. Los detallistas consideraban esencial que pudieran vender la línea completa. El demandado requería que los detallistas aceptaran no manejar líneas rivales. En principio, los fabricantes rivales podrían crear sus propias tiendas de menudeo, pero ¿quién comprará en ellas si no puede encontrarse allí la marca más popular? Los fabricantes rivales tendrían que crear una línea tan extensa y popular como la línea de Standard Fashion, y eso sería difícil de lograr.

La venta exclusiva, el nombre de la forma contractual utilizada por Standard Fashion, es de ordinario una práctica eficiente.[58] Estimula al vendedor para promover la marca del fabricante; en Standard Fashion disminuía también el peligro de la piratería de estilos. Estas preocupaciones se relacionan estrechamente entre sí. En ausencia de la distribución exclusiva, el vendedor podría verse tentado a usar los patrones de Standard Fashion para atraer clientes a su tienda y desviarlos luego hacia marcas más baratas que se asemejaran a los patrones (quizá sean copias de ellos), mediante un precio de mayoreo muy bajo para tales marcas. El precio sería bajo porque el fabricante de los rivales baratos, aprovechándose gratuitamente de la inversión de Standard Fashion para crear patrones de diseño populares, tendría costos menores que los de Standard Fashion.

Lo que distinguía al caso de Standard Fashion de un caso de distribución exclusiva común y corriente era la existencia de economías de escala al nivel de la distribución, que son análogas a las economías de escala en el consumo. Los consumidores no deseaban andar errando de una tienda a otra

[56] Véanse, por ejemplo, Stanley M. Besen y Joseph Farrell, "Choosing How to Compete: Strategies and Tactics in Standardization", *J. Econ. Perspectives*, primavera de 1994, 117; Joseph Farrell y Garth Saloner, "Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation", 76 *Am. Econ. Rev.,* 940 (1986); Michael L. Katz y Carl Shapiro, "Technology Adoption in the Presence of Network Externalities", 94 *J. Pol. Econ.,* 822 (1986); véase también después XXI.4. Una fuente estrechamente relacionada de la dependencia de la ruta es la de los "costos del cambio": habiendo aprendido a usar una computadora, un consumidor podría resistirse al cambio a un tipo de computadora diferente, aunque fuera superior, debido al costo de aprender a operar un tipo diferente.

[57] Standard Fashion Co. *vs.* Magrane-Houston Co., 258 U.S., 346 (1922).

[58] Véase Howard P. Marvel, "Exclusive Dealing", 25 *J. Law Econ.,* 1 (1982).

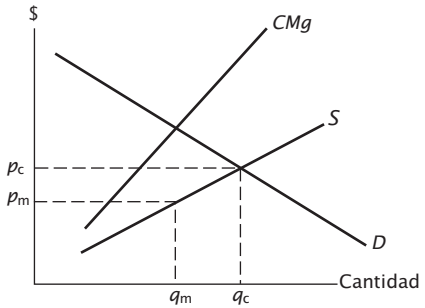
para conseguir patrones de vestidos. Querían una línea completa en cada tienda. Por lo tanto, cualquiera que ingresara al negocio de los patrones de vestidos tenía que proporcionar una línea completa si se le excluía de las tiendas que operaban con la línea de la empresa dominante en ese campo. (Podemos imaginar una tienda de departamentos reuniendo los diseños de varios diseñadores de patrones de vestidos femeninos, cada uno de los cuales se especializara en un patrón, en una línea completa para competir con la línea completa de Standard Fashion, pero el caso de Standard Fashion se falló en 1922 y muchas ciudades no tenían tiendas de departamentos en aquellos días.) No hay duda de que la restricción de sus distribuidores de menudeo le costaba algo a Standard Fashion, pero es posible que ese costo fuese menor que el aumento de sus ganancias monopólicas esperadas de la exclusión de nuevas entradas al obligar a los neófitos potenciales a entrar con una línea completa —un proyecto arriesgado, como si una empresa no pudiera fabricar aviones comerciales sin fabricar también aviones militares—. Standard Fashion no podía extraer dos veces un precio monopólico de sus clientes, cobrándoles primero lo que el mercado soportara y obligándolos luego a celebrar contratos de distribución exclusiva. Pero es posible que haya podido alargar la duración de su monopolio.

Debe ser obvia la analogía con una externalidad de red de la nueva economía. Una empresa podría desear entrar al mercado produciendo un componente de la red o un servicio de valor agregado, pero, si un competidor que es propietario de la red o tiene un contrato de distribución exclusiva con ella se niega a cooperar con la empresa, ésta tendrá que duplicar la red a fin de obtener la distribución de su producto.

X.12. Boicots; origen del monopsonio

En el caso de Eastern States, un grupo de distribuidores de madera al menudeo aceptaron no comprar madera de ningún mayorista que se dedicara al negocio de la madera al menudeo en competencia con un miembro del grupo de minoristas.[59] Pero el acuerdo podría disuadir a los mayoristas de entrar a ese negocio sólo si los minoristas tuvieran poder monopsónico frente a los mayoristas. La gráfica X.3 representa la demanda derivada de algún insumo por parte de una empresa (es decir, derivada de la demanda del producto final de la empresa). Se supone que el precio de oferta se relaciona positivamente con la cantidad del insumo comprada, lo que implica que la producción de éste involucra el uso de recursos especializados. Así, el costo marginal del insumo aumenta más deprisa que el precio de oferta (véase X.9 anterior) —es decir, cada incremento del precio aumenta los costos totales del insumo para la

[59] Eastern States Retail Lumber Dealers' Ass'n. *vs.* United States, 234 U.S., 600 (1914).



GRÁFICA X.3

empresa en una cantidad mayor— porque el aumento del precio se aplica no sólo a la compra incremental sino a todas las compras anteriores del insumo (suponiendo que no hay discriminación de precios). (Así pues, el costo marginal es al precio de oferta como el ingreso marginal es al precio de venta.)

Las ganancias del comprador se maximizan cuando limita sus compras del insumo al nivel, q_m en la gráfica X.3, en que el costo marginal es igual al valor marginal. A este nivel, el precio es menor que si la competencia en el mercado de compra condujera a los compradores a eliminar las ganancias monopsónicas generadas en q_m, elevando el precio hasta p_c.

La atracción de la monopsonización para el monopsonista depende de la pendiente positiva de la curva de oferta. Pues si esta curva fuera horizontal la reducción de la cantidad del insumo comprado no generaría ganancias monopsónicas, porque el precio del insumo sería el mismo con cualquier cantidad que se comprara.

La curva de la oferta es positiva sólo cuando un insumo consume recursos que serían menos valiosos en otros usos. Una vez que se ha construido una vía de ferrocarril, su valor en usos alternativos queda severamente limitado, de modo que un monopsonista de servicios de ferrocarril podría ser capaz de limitar el precio que paga por tales servicios a un nivel tal que el ferrocarril no podría recuperar su inversión en las vías.[60] Una mina de carbón es otro buen ejemplo de un recurso que no puede ser reubicado sin una pérdida de valor enorme.[61]

Si los minoristas de madera en el caso de Eastern States tenían poder monopsónico, ¿por qué no lo ejercieron directamente reduciendo la cantidad de madera que compraban a los mayoristas? Quizá les resultara más fácil a los minoristas ponerse de acuerdo para eliminar a los mayoristas rivales que ponerse de acuerdo sobre el precio monopsónico óptimo de la madera. Otra

[60] Esa inversión se aproxima en la gráfica X.3 —suponiendo que el ferrocarril no utiliza otros recursos especializados— por el área triangular limitada por el eje vertical, p_c, y la curva de oferta.

[61] Véase el problema 5 al final del capítulo XIII.

posibilidad es que los mayoristas fuesen menos eficientes que los minoristas en la distribución de menudeo y se viesen atraídos al mercado de menudeo sólo porque los minoristas estaban cobrando un precio de cártel. Entonces el boicot aumentaría la eficiencia del cártel. Los precios monopsónicos no lo harían.

X.13. ¿Existe el monopsonio benigno?

Se ha argumentado que, en un esfuerzo por controlar los costos de la atención a la salud (cerca de la mitad de los cuales los asume el gobierno federal), el gobierno debería utilizar su enorme "poder de compra" (es decir, su poder monopsónico) para reducir los precios de los medicamentos. El gobierno es un gran comprador indirecto de medicamentos ya que financia los programas Medicare y (con los estados) Medicaid, además de brindar atención médica al personal militar y a los veteranos y seguro médico a los empleados civiles del gobierno. Dado que el gobierno compra, o (más comúnmente) controla (actuando en realidad como un comprador intermedio) la compra de una fracción muy grande de las ventas totales de medicamentos, podría reducir en cierta cantidad arbitraria los precios de los medicamentos por los que está dispuesto a hacer un rembolso y, sin embargo, las compañías farmacéuticas continuarían vendiéndole.

Sin embargo, las compañías farmacéuticas reducirían su producción, pero, dada la estructura de los costos de los medicamentos, no sería de inmediato. El costo de desarrollar y de obtener la aprobación de la FDA para un medicamento nuevo representa un gran porcentaje del costo total y se incurre en él antes de que el medicamento salga al mercado. Por lo general, el costo de producir el medicamento es muy bajo. Por lo tanto, el precio de venta se debe principalmente a los costos de desarrollo de preventa, y, por ello, el efecto de la negativa del gobierno a pagar el precio actual consistiría, al reducir las ganancias de la empresa (en especial sus ganancias proyectadas de los medicamentos nuevos), en determinar una reducción en el desarrollo de medicamentos nuevos antes que la producción de los medicamentos existentes. Este análisis no se aplica a los medicamentos genéricos —no patentados—, que se venden de ordinario a un precio cercano al costo marginal. El análisis de monopsonio común se aplicaría a aquellos medicamentos: el gobierno pagaría un precio que obligara al productor a bajar su curva de la oferta, reduciendo la producción.

Frenar el desarrollo de medicamentos nuevos sería políticamente atractivo porque la reducción de los costos de los medicamentos financiados por el gobierno se experimentaría de inmediato, pero el costo de un desarrollo más lento (porque serían menos rentables) de medicamentos nuevos se aplazaría, y de hecho sería invisible en gran medida porque nadie sabría a ciencia cierta cuánto más rápidos habrían sido el desarrollo y la introducción de medica-

mentos nuevos de no haber sido porque el gobierno ejerció su poder monopsónico.

X.14. Los daños antimonopólicos

Aunque los remedios equitativos, tales como las prohibiciones y la privación, y los remedios penales (multas y prisión) se utilizan también para hacer cumplir las leyes antimonopólicas federales, el remedio de los daños plantea los problemas económicos más fascinantes y es aquí donde concentraremos nuestro análisis.

El demandante exitoso en un juicio antimonopólico tiene derecho a una triplicación de sus daños compensatorios, de modo que dos terceras partes de las compensaciones antimonopólicas otorgadas por daños representan daños punitivos. En el capítulo VII vimos que esto sería sensato si hubiera una probabilidad de un tercio de aprehender al violador antimonopólico. En efecto, sin embargo, la probabilidad de las violaciones antimonopólicas cometidas abiertamente (las fusiones, por ejemplo) se aproxima a uno, mientras que la probabilidad para las violaciones ocultas (principalmente las conspiraciones para la fijación de los precios), aunque menor a uno, no es siempre igual a un tercio.

¿Cómo debieran computarse los daños compensatorios? En términos de la gráfica IX.5, ¿debieran ser tales daños iguales a *PM* (la pérdida de peso muerto)?, ¿*MP* (el precio del cártel u otro precio monopólico menos el precio competitivo multiplicado por la producción del monopolio)?, ¿*PM* + *MP*? ¿Más? *PM* sólo mide una parte del costo impuesto por el monopolio: el costo que soportan quienes dejan de comprar el producto cuando el precio aumenta desde el nivel competitivo hasta el monopólico. No considera el costo para los consumidores que continúan comprando el producto pero pagan más. Su costo es *MP,* y el costo social total es por lo tanto *MP* + *PM*. Es cierto que *MP* no es un costo social neto, sino sólo un pago de transferencia, por lo menos si omitimos la tendencia de las ganancias monopólicas a transformarse en costos de la obtención o la defensa del monopolio. Sin embargo, queremos que el monopolista, al decidir si le convendrá monopolizar o no, compare su ganancia con la pérdida de todos los perjudicados por el monopolio, lo que incluye a los consumidores que continúan comprando y a los que se cambiaron a sustitutos (inferiores). Para que haga esto debemos fijar daños iguales a los costos sociales totales, no a los netos, del monopolio.

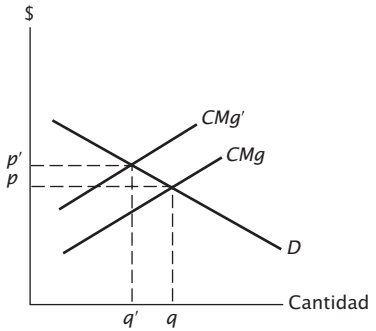
Tampoco sería suficiente hacer los daños iguales a *MP,* según la teoría de que eso eliminaría las ganancias del monopolio, el que en consecuencia dejaría de ser atractivo. Supongamos que el monopolio generaría algún ahorro pequeño de los costos del monopolista. Entonces *MP* —por lo menos si se calcula de acuerdo con la curva de costo anterior al monopolio, es decir, de acuerdo con lo que los consumidores han perdido y no con lo que el monopolista ha ganado— sería menor a la ganancia obtenida del monopolio por el

monopolista, de modo que éste no sería disuadido. Y, aun así, si el ahorro de costos fuese menor que *PM,* desearíamos impedirlo; el monopolio no estaría justificado por sus costos.

Sin embargo, la posibilidad de ahorrar costos demuestra que sería inconveniente fijar los daños (dejando de lado las posibilidades del ocultamiento) por arriba de la suma de *MP* y *PM*. Pues eso disuadiría a los monopolios que generaran ahorros de costos mayores a la pérdida de peso muerto… si suponemos, realistamente, que la ley antimonopólica no excusa de responsabilidad a todos ellos.

¿Cuáles serían los daños si forman un cártel las empresas que no controlan todo el mercado? La discusión anterior estableció que el cártel podría ser capaz todavía de cobrar un precio monopólico, aunque no tan alto como si controlara todo el mercado. Supongamos que logra elevar el precio 10%. Se beneficiarán las empresas que estén fuera del cártel; es probable que vendan al mismo precio que los cartelistas, ya que si empiezan a bajar los precios el cártel podría derrumbarse y es posible que todos los vendedores estén en una situación peor. Pero los consumidores de los vendedores que no pertenecen al cártel perderán junto con los consumidores de los cartelistas. ¿Debieran los miembros del cártel ser responsables no sólo por sus propias ganancias monopólicas sino también por las de los vendedores inocentes (quienes no han violado ninguna ley)? No debieran serlo. Es cierto que los clientes de tales vendedores habrán sufrido a resultas del cártel. Pero si han de cargarse tales pérdidas al cártel, también debieran acreditársele los beneficios (ganancias) que confirió a los vendedores inocentes; y las dos cantidades quedan empatadas. Ésta es una aplicación directa del análisis de las pérdidas "económicas" referidas en los capítulos VI y VIII: las pérdidas contrarrestadas por las ganancias de otros no debieran contarse en el monto de los daños, aunque las pérdidas fueran causadas por el cuasidelito. Con todo, una dificultad es que las empresas de la franja tienden a tener costos más altos que los cartelistas; se supone que es por eso que son pequeñas, una franja. No obstante, esto significa que parte del excedente del consumidor de que se apropian estas empresas se transforma en un costo en vez de un excedente del productor. Este costo es una pérdida de peso muerto por la cual, en principio, se debe responsabilizar a los cartelistas.

A menudo, los consumidores no compran directamente a los fabricantes sino a intermediarios. Supongamos que un monopolista de zapatos vende a 10 distribuidores que por su parte revenden a un millón de minoristas, que a su vez revenden a un millón de consumidores. Se debiera permitir que los 10 distribuidores demanden al monopolista por todo el recargo monopólico, aunque con toda probabilidad habrán trasladado el grueso del recargo a los detallistas, que a su vez lo habrán trasladado a los consumidores. El monto del traslado depende de la elasticidad de la demanda. Entre más elástica sea la demanda, menor será la fracción del recargo que se traslade a la siguiente etapa de la distribución. El proceso del traslado se representa en la gráfica X.4.



GRÁFICA X.4

La curva con prima incluye el recargo, la curva sin prima muestra los costos antes de que se imponga el recargo y la distancia vertical entre las dos representa el monto del recargo por unidad de insumo comprado.

Dependiendo del grado del traslado, el juicio de los distribuidores podría rendirles ganancias súbitas; pero lo más importante —la desistencia del monopolio— se habrá logrado con mayor eficacia que si se prohibieran tales juicios. Y es posible que no haya ninguna ganancia súbita. Dado que está bien establecida la regla de que no hay defensa del traslado, como lo está ahora, se puede esperar que los intermediarios cobren precios menores en general. Los costos netos de sus compras al fabricante serán menores, porque un intermediario tendrá una ganancia esperada del hecho de poder demandar si el precio resulta ser un precio monopólico ilegal. Los precios menores cobrados por los intermediarios compensarán a los consumidores de perder su propio, menos valioso derecho a demandar. Su derecho es menos valioso porque, encontrándose más alejados del fabricante y siendo más numerosos, los consumidores son aplicadores de los remedios antimonopólicos menos eficientes. Si esto es correcto, ¡los consumidores han pasado en realidad a una situación mejor por el rechazo de la defensa del traslado!

Pero ¿debiera preocuparnos que el tener derecho a una triplicación de los daños hiciera que los compradores esperaran demasiado antes de demandar, prolongando así el periodo de los precios monopólicos?[62] Supongamos que el precio competitivo de cierto bien es 10 dólares y el precio monopólico es 15 dólares. Por lo tanto, los daños simples son 5 dólares y los daños triples son 15. Esto significa que cada compra del bien le cuesta al comprador… ¡5 dólares! Por lo tanto, tiene un incentivo (limitado sin embargo por la ley de limitaciones y por su tasa de descuento) para demorar el juicio indefinida-

[62] Véase Pablo T. Spiller, "Treble Damages and Optimal Suing Time", 9 *Research in Law & Econ.,* 45 (1986).

mente. Porque, sabedor de este incentivo, el monopolista advertirá que los precios monopólicos prolongados podrían conducir a daños astronómicos.

Una interrogante importante acerca de los daños antimonopólicos es si debiera permitirse que los competidores y, asimismo, los compradores demandaran el pago de daños. A primera vista, la respuesta pareciera ser negativa si el propósito de la ley antimonopólica es promover la asignación eficiente de los recursos. El monopolio crea ineficiencia al meter una cuña entre el costo de oportunidad (= precio competitivo) y el precio (monopólico); el bienestar de los competidores no está ni aquí ni allá. Pero a veces un competidor es un aplicador más eficiente de la ley antimonopólica que un consumidor porque, como un intermediario, posee más información y es probable que tenga más en juego en la aplicación de la ley antimonopólica. Consideremos los precios depredadores. Los consumidores se benefician a corto plazo; es sólo después de que la depredación triunfa y el monopolista eleva su precio que los consumidores empiezan a verse perjudicados. Por lo tanto, es improbable que hubiera muchos juicios iniciados por los consumidores contra los precios depredadores. Son los competidores quienes se ven perjudicados a corto plazo y, por ende, es probable que sean quienes demanden.

Pero *¿cuál* debiera ser exactamente el motivo por el que se permitiría demandar a los competidores perjudicados? El perjuicio que sufren no guarda ninguna relación necesaria con los costos sociales de la depredación (que son los costos sociales del monopolio en el periodo posterior a la depredación, descontados al valor presente). Pero si las víctimas de los precios depredadores pueden contar con ser indemnizadas por el perjuicio que les causa la depredación, el depredador no podrá intimidarlas con la amenaza de los precios depredadores.

En el caso Brunswick,[63] la Corte Suprema otorgó una limitación importante sobre los juicios de daños iniciados por los competidores. A demandó a B por haber adquirido a C, un competidor de A, y reclamó daños basados en el perjuicio que C le había causado a A gracias a su fortalecimiento a manos de B: si no hubiese ocurrido la adquisición, C podría haber abandonado la industria y, en consecuencia, habrían sido mayores la producción y las ganancias de A. Se suponía que la adquisición había sido ilegal porque había creado un peligro (que sin embargo no se había materializado) de que B financiara los precios depredadores de C. Con todo, no se permitió que A obtuviera el pago de daños que demandaba. Aunque estos daños eran causados por una adquisición ilegal, tenían la misma significación económica que el daño causado a los consumidores por el precio de cártel de empresas que no están en el cártel: eran daños compensados en otra parte del sistema económico. Si la adquisición de B por parte de C le permitiera a C continuar compitiendo con A sin recurrir a tácticas impropias tales como los precios depre-

[63] Brunswick Corp. *vs.* Pueblo Bowl-O-Mat, Inc., 429 U.S., 477 (1977).

dadores, toda pérdida de A debe haberse compensado con creces por los beneficios de esta competencia para los consumidores; porque sólo haciéndoles ofertas que fuesen mejores (aunque por hipótesis eficientes, no por debajo del costo) a los consumidores podría C habérselos arrebatado a A. Habría una disuasión excesiva si se obligara a B a pagarle daños a A. Compárese con VI.7 anterior.

Lecturas recomendadas

Baird, Douglas G., Robert H. Gertner y Randal C. Picker, *Game Theory and the Law,* 57-63, 165-187 (1994).

Carlton, Dennis W., y Jeffrey M. Perloff, *Modern Industrial Organization* (4a ed., 2005).

Crane, Daniel A., "Chicago, Post-Chicago, and Neo-Chicago", 76 *U. Chi. L. Rev.*, 1911 (2009).

Dick, Andrew R., "Learning by Doing and Dumping in the Semiconductor Industry", 34 *J. Law & Econ.,* 133 (1991).

Economics of Antitrust Law (2 vols.; Benjamin Klein y Andres V. Lerner [comps.], 2008).

Hemphill, C. Scott, y Tim Wu, "Parallel Exclusion", 122 *Yale L. J.*, 1182 (2013).

Hovenkamp, Herbert, *The Antitrust Enterprise: Principle and Execution* (2005).

Kaplow, Louis, "Competition Policy and Price Fixing" (2013).

Klein, Benjamin, "Market Power in Antitrust: Economic Analysis After *Kodak"*, 3 *Supreme Ct. Econ. Rev.,* 43 (1993).

———, "Exclusive Dealing as Competition for Distribution 'On the Merits'", 12 *Geo. Mason L. Rev.*, 119 (2003).

Landes, William M., "Optimal Sanctions for Antitrust Violations", 50 *U. Chi. L. Rev.,* 652 (1983).

———, y Richard A. Posner, "Should Indirect Purchasers Have Standing to Sue under the Antitrust Laws: An Economic Analysis of the Rule of *Illinois Brick"*, 46 *U. Chi. L. Rev.,* 602 (1979).

Lopatka, John E., "Solving the Oligopoly Problem: Turner's Try", 41 *Antitrust Bull.,* 843 (1996).

Page, William H., "Antitrust Damages and Economic Efficiency: An Approach to Antitrust Injury", 47 *U. Chi. L. Rev.,* 467 (1980).

Posner, Richard A., *Antitrust Law* (2a ed., 2001).

Rasmusen, Eric B., J. Mark Ramseyer y John S. Wiley Jr., "Naked Exclusion", 81 *Am. Econ. Rev.,* 1137 (1991).

Salop, Steven C., "Exclusionary Conduct, Effect on Consumers, and the Flawed Profit-Sacrifice Standard", 73 *Antitrust L. J.*, 311 (2006).

Stigler, George J., *The Organization of Industry*, caps. 5-10, 21 (1968).

Problemas

1. Enseguida aparece una lista de ciertos casos de antimonopolio destacados que falló la Corte Suprema. Lea cada decisión. Luego conteste las siguientes preguntas: ¿Por qué el demandado o los demandados adoptaron en efecto la práctica cuestionada: para monopolizar o para algún otro propósito? ¿Hasta qué punto descansa la decisión de la Corte en una consideración de las consecuencias económicas de la práctica cuestionada? ¿Emplea la Corte correctamente el análisis económico?

United States *vs.* Reading Co., 253 U.S., 26 (1920).
Chicago Board of Trade *vs.* United States, 246 U.S., 231 (1918).
United States *vs.* Container Corp. of America, 393 U.S., 333 (1969).
United States *vs.* Sealy, Inc., 388 U.S., 350 (1967).
United States *vs.* Continental Can Co., 378 U.S., 441 (1964).
United States *vs.* Arnold, Schwinn & Co., 388 U.S., 365 (1967).
United States *vs.* Singer Mfg. Co., 374 U.S., 174 (1963).
Sugar Institute, Inc. *vs.* United States, 297 U.S., 553 (1936).
Arizona *vs.* Maricopa County Medical Society, 457 U.S., 332 (1982).
United States *vs.* United States Gypsum Co., 438 U.S., 422 (1978).
National Collegiate Athletic Association *vs.* Board of Regents, 486 U.S., 85 (1984).
California Dental Association *vs.* FTC, 526 U.S., 756 (1999).
Aspen Skiing Co., *vs*. Aspen Highlands Skiing Corp., 472 U.S., 585 (1985).

2. Con frecuencia, las aerolíneas cobran más por un boleto sencillo que por un boleto de viaje redondo. ¿Podría ser ésta una práctica de exclusión?

3. ¿Puede la teoría del cártel ayudarle a contestar la pregunta siguiente: en una alianza militar, tal como la OTAN, ¿comprometerán las naciones más pequeñas recursos mayores, menores o iguales a la alianza, en proporción a su tamaño? Si se le dificulta, consulte Todd Sandler y Keith Hartley, "Economics of Alliances: The Lessons for Collective Action", 39 *J. Econ. Lit*., 869 (2001).

4. Analice la proposición siguiente: La política antimonopólica es profundamente contraria al logro de las cantidades económicamente óptimas de: *a)* protección ambiental y *b)* información y seguridad del producto para el consumidor.

5. El Departamento de Justicia ofrece inmunidad judicial al primer miembro de un cártel que confiese sobre su actividad en el cártel. ¿Podría un programa de este tipo alentar o desalentar la cartelización? ¿Podría hacer ambas cosas? Véase Nathan H. Miller, "Strategic Leniency and Cartel Enforcement", 99 *Am. Econ. Rev.*, 750 (2009).

6. ¿Debieran las empresas sin fines de lucro, como las universidades privadas y los hospitales, estar exentas de la ley antimonopólica? Véase Tomas J. Philipson y Richard A. Posner, "Antitrust in the Not-for-Profit Sector", 52 *J. Law and Econ.*, 1 (2009).

XI. LA PROPIEDAD INTELECTUAL

XI.1. ¿Las patentes son una propiedad?

Las ediciones anteriores de este libro introdujeron discusiones breves de la propiedad intelectual en el capítulo III (La propiedad). El campo de la propiedad intelectual se ha vuelto tan grande, importante y controvertido que requiere un capítulo aparte. Además, aunque existen afinidades obvias y destacadas entre la propiedad tangible y la propiedad intelectual, las diferencias son tan grandes que el mismo término "propiedad intelectual" disfraza algunas de las razones de peso para no considerar los inventos, nombres comerciales, secretos comerciales y medios de expresión, tales como las películas y novelas, como una propiedad en absoluto, sino como frutos *sui generis* de investigación y creación mental, que en ocasiones requieren protección legal, pero una protección legal distinta de la que se concede a la propiedad tangible.

Desde luego, existen afinidades entre las dos formas de la "propiedad". Por ejemplo, aun cuando nada parecería estar más alejado de los tesoros hundidos que los inventos patentados, uno de los problemas económicos creados por el otorgamiento de patentes se asemeja a un problema económico con una propiedad abandonada. Las ideas se crean en cierto sentido, pero en otro sentido se encuentran. Supongamos que, quienquiera que invente determinado artefacto, si se le permite excluir a otros de su uso mediante el otorgamiento de una patente, podrá vender la patente a un fabricante por un millón de dólares; y que el costo esperado de la invención del artefacto es de 250 000 dólares. Varios inventores podrían tratar de ser los primeros en inventarlo. Su competencia (una "carrera de patentes") hará que el invento aparezca más pronto. Eso es algo bueno… pero supongamos que se inventa apenas un día antes de lo que lo habría hecho si sólo hubiera habido un inventor. El valor de tener el adminículo un día antes será menor que el costo de un aumento de cuatro veces el costo del invento.

Esto implica que la *negación* de un derecho de propiedad, incluido un derecho de propiedad intelectual, pueda ser tanto un instrumento economizador como la creación de uno. Como un ejemplo adicional, adviértase que la protección legal de una marca comercial depende de que el propietario de la marca comercial esté vendiendo efectivamente el producto o el servicio designado por dicha marca. No se pueden simplemente soñar nombres para productos que usted o alguien más podrían desear vender algún día, y registrarlos en la Oficina de Patentes y Marcas Registradas (USPTO o PTO) para, al realizar

esta acción, obtener el derecho a excluir a otros del uso de estos nombres. Permitir tal "jaloneo" de marcas comerciales atraería recursos excesivos a la actividad de pensar marcas comerciales, y también saturaría el registro con millones de marcas, haciendo que fuera más costoso para los vendedores buscar el registro a fin de evitar la violación de una marca registrada. (El desordenado abarrotamiento del registro es un gran problema con las patentes: en la actualidad existen alrededor de dos millones de patentes vigentes en los Estados Unidos, cualquiera de las cuales podría ser la base de una demanda por violación de patente.)

Incluso existe un paralelo entre la propiedad intelectual y las cláusulas restrictivas. Según la doctrina de Europa continental del "derecho moral", la cual se ha abierto camino en el derecho norteamericano (en parte debido a los compromisos de los Estados Unidos derivados del Convenio de Berna sobre los *copyrights*), un artista tiene (sin necesidad de reservarlo de manera explícita) el derecho de impedir la mutilación, destrucción o atribución errónea de sus pinturas u otras obras de arte incluso después de haber vendido la obra, tal vez a un distribuidor que luego la revendió a un coleccionista o a un museo. Además, el derecho sobrevive a la muerte del artista, convirtiéndose en un activo de sus herederos. De este modo, el derecho del artista a la integridad de su arte irá aparejado a la obra, al igual que un usufructo de la tierra. La justificación económica es que el valor de la obra futura o no vendida del artista, así como sus obras pertenecientes a otros compradores, puede disminuir por actos que afecten el valor de una de sus obras en particular.[1]

Se tiene que permitir que el derecho moral sobreviva a la muerte del artista, con el fin de evitar un problema de congestión similar a aquel que involucra un pastizal natural (véase antes III.1). En principio, los propietarios de sus otras obras debieran ser los responsables de hacer cumplir este derecho póstumo porque son las personas perjudicadas por su violación, pero existen economías en el hecho de dar a los herederos del artista el derecho exclusivo a demandar.

Una diferencia jurídica interesante entre la propiedad física y la propiedad intelectual es que esta última, una vez que se abandona, nunca se puede reapropiar. A diferencia de la propiedad física (véase III.13 anterior), ésta entra en el dominio público, donde permanece libremente utilizable —para siempre— por cualquier persona. La diferencia en el tratamiento tiene sentido económico. La propiedad intelectual es un ejemplo de lo que los economistas llaman un "bien público", lo que significa que es un bien que puede consumirse sin reducir su consumo por parte de cualquier otra persona. Es mucho más probable que la propiedad física sea un bien privado. Que una persona

[1] Henry Hansmann y Marina Santilli, "Authors' and Artists' Moral Rights: A Comparative Legal and Economic Analysis", 26 *J. Legal Stud.*, 95 (1997). Véase un punto de vista escéptico sobre los derechos morales en William M. Landes y Richard A. Posner, *The Economic Structure of Intellectual Property Law*, cap. 10 (2003).

lea un libro no impide que otra persona lea el mismo libro en un ejemplar distinto, pero que una persona coma un sándwich impide que cualquier otra persona coma ese sándwich. Excepto por los usos especiales, como los parques públicos, los terrenos y otras propiedades físicas, no se pueden inscribir en el dominio público (o dejarse librados al uso en este ámbito si se abandonan) porque el uso compartido es ineficiente. Queremos a alguien que pueda reapropiárselos si han sido abandonados, para que pueda explotar la propiedad a fin de crear valor.

XI.2. PATENTES E INNOVACIÓN[2]

El aspecto más controvertido de la ley de propiedad intelectual hoy en día es la ley de patentes. Para entender las controversias, tenemos que entender la economía de la invención.

Supongamos que cuesta 10 millones de dólares inventar un nuevo tipo de batidora de alimentos. Una vez inventada, el costo marginal de producir y vender cada batidora es de, digamos, 50 dólares (los 10 millones no son un costo marginal, sino un costo fijo, porque no varía en función de la producción), y la demanda se estima en un millón de batidoras si el precio es de 60 dólares por cada una. A menos que el fabricante pueda cobrar 60 dólares por cada batidora, no recuperará sus costos de invención (10 millones) más los de producción (50 millones).[3] Pero si otros fabricantes afrontan el mismo costo marginal, la competencia (en ausencia de patentes) bajará el precio a 50 dólares, el esfuerzo de recuperación disminuirá y, previendo esto, el fabricante no hará la invención en primer lugar: no sembrará si no es capaz de cosechar.

Además, en un mundo sin patentes, la actividad de invención estaría sesgada hacia los inventos que pudieran mantenerse en secreto, del mismo modo que la ausencia de derechos de propiedad sobre bienes físicos sesgaría la producción hacia las cosas que implican una inversión preparatoria mínima, como vimos antes en III.1. (Sin embargo, algunos inventores prefieren el secreto a las patentes como el modo de protegerse contra los imitadores. Véase después XI.3.)

La teoría de la innovación óptima tiene implicaciones importantes para el derecho de patentes. Considérese la patentabilidad de los inventos que mejoran una invención patentada existente pero que no están autorizados por el titular de la patente. Aunque una empresa tiene un incentivo para autorizar a los inventores independientes de mejoras en sus patentes, su capacidad para identificar a los independientes más prometedores estará limitada por sus propios conocimientos, experiencias y cultura como empresa. Por lo tanto, es

[2] Véase en general John F. Duffy y Robert Patrick Merges, *Patent Law and Policy: Cases and Materials* (5ª ed., 2011).

[3] Para más fácil, omito el precio y, por lo tanto, los ingresos para los distintos niveles de producción.

valioso alentar a las empresas a investigar y desarrollar tecnologías dominadas por una patente o patentes en poder de otra empresa, como lo hace la ley al permitir que las mejoras aportadas por esas empresas se patenten sin la autorización de la empresa dominante. Sin embargo, ésta es la segunda mejor solución, ya que crea un monopolio bilateral: la empresa dominante no puede incorporar la mejora a su producto sin la autorización del titular de la patente de mejora, y quien hace la mejora no puede vender el producto mejorado sin autorización de la empresa dominante.

Es posible otorgar una patente antes de que exista una producción comercial o incluso un prototipo de la invención. Sólo se necesita un diagrama o descripción que permita a una persona experta crear el producto o proceso patentado. El establecimiento anticipado de los derechos de propiedad fomenta la inversión,[4] como sabemos por la discusión de los derechos de propiedad en el capítulo III. Pero, al tratar de reducir su invención a la práctica, el inventor puede no seguir el camino óptimo de desarrollo. Al otorgarle el derecho exclusivo y reducir así la cantidad de inventos paralelos, la ley podría, por lo tanto, estar retrasando el desarrollo tecnológico. Y nótese la tensión entre permitir la patente anticipada y permitir las patentes de mejora no autorizadas.

Como dije, actualmente hay cerca de dos millones de patentes vigentes en los Estados Unidos, lo que aumenta los costos de búsqueda de patentes y de violación involuntaria y complica su tramitación (el proceso de convencer a la Oficina de Patentes y Marcas Registradas de los Estados Unidos de otorgar una patente) y el litigio (al aumentar el número de patentes que se pueden citar como "avances previos" se invalida la patente en disputa). Elevar los parámetros para la concesión de patentes, acortar los plazos de las patentes[5] o incluso negar la patentabilidad de los inventos en industrias específicas —todas ellas medidas imaginables para reducir el exceso de patentes— no pondría fin a la actividad de invención (incluso cuando un invento no pudiera mantenerse en secreto) y de hecho podría estimularla mediante la reducción de los costos y los riesgos que enfrentan los inventores por las marañas de patentes. Fabricar algunos inventos no es tan costoso de modo que se tengan que recuperar los costos durante un largo periodo de tiempo; el inventor podría ser capaz de recuperarlos antes de que los competidores puedan adaptarse

[4] Edmund W. Kitch, "The Nature and Function of the Patent System", 20 *J. Law & Econ.,* 265 (1977).

[5] Lo cual se puede hacer de forma indirecta al aumentar las cuotas para mantener las patentes. Véase James Bessen y Brian J. Love, "Make the Patent 'Polluters' Pay: Using Pigovian Fees to Curb Patent Abuse" (Boston U. School of Law, Law and Econ. Research Paper Series, núm. 13-23 (2013)). En la actualidad, para mantener una patente hasta su término, las cuotas deben pagarse tres y medio años después de la concesión de la patente, luego siete y medio años, después once y medio años. Las cuotas son razonables, van desde 400 dólares, la cuota de tres y medio años para una "microentidad", hasta 7 400 dólares, la cuota de once y medio años para una entidad mayor.

para duplicar el invento y vender por debajo del precio del productor original. O el costo de conseguir la invención puede ser bajo en relación con el costo de producción de los bienes que componen la invención, por lo que la ventaja en costos de un imitador no será enorme. O el público podría identificar la marca de la primera empresa con el producto y seguir comprándolo incluso después de que haya competencia, y a un precio mayor. Y muchos productos nuevos sólo tienen una vida esperada breve, de modo que 20 años de protección de patentes no otorgan ningún beneficio real al productor, excepto para permitirle obtener los derechos de licencia de las empresas que quieren hacer un producto diferente que incluya su invención. En ocasiones el tiempo de elaboración del inventor, antes de que los competidores reproduzcan el producto que constituye la invención, aunque sea poco, es suficiente para darle una ventaja en la "curva de aprendizaje" sobre los competidores. La curva de aprendizaje es una curva con pendiente negativa que representa el costo de producción como una función de la experiencia en la producción. Cuanta más experiencia tenga una empresa en la producción de un producto, menor será su costo, porque la experiencia prueba formas de reducir costos y mejorar el producto. La existencia de la curva de aprendizaje pone en desventaja a los rezagados al competir con el productor que llegó antes (el "primer motor", como se le llama). Y, por último, los costos de reproducción pueden ser muy altos, en especial si el inventor puede mantener su invención en secreto.

Así, es probable que una gran cantidad de patentes no tengan efecto positivo alguno en la invención; gran parte de esas patentes tienen por objeto monopolizar un producto o simplemente son defensivas. Y patentar plantea una serie de problemas económicos. Las patentes fomentan la búsqueda dispendiosa de rentas (carrera de patentes); favorecen la inversión en los tipos de actividad de invención que generan patentes (si hay dos actividades igualmente valiosas que son sustitutas y sólo en una de ellas el inventor puede obtener un derecho de propiedad sobre los bienes o servicios que produce la actividad, la inversión se inclinará a favor de esa actividad, lo que generará una inversión insuficiente en la otra); obstaculizan la rivalidad en la actividad inventiva, y, al abrir una brecha entre el precio y el costo (al gozar como lo hace de un monopolio, el dueño de la patente no se ve obligado a establecer precios a nivel del costo marginal), desvían a algunos clientes potenciales del producto patentado hacia sustitutos menos eficientes. Una vez que se hace una invención, sus costos son irrecuperables, están hundidos; en términos económicos, son iguales a cero. Por lo tanto, un precio que incluye una regalía para el inventor excederá el costo de oportunidad del producto en el que queda encarnada la invención. Sin embargo, analíticamente esta brecha es igual al costo de una cerca para demarcar un derecho de propiedad sobre la tierra; es un costo indispensable por utilizar el sistema de derechos de propiedad para la asignación de recursos.

La ley utiliza varios instrumentos para tratar de minimizar estos problemas:

1. Una patente expira después de 20 años (existen algunas excepciones), en lugar de ser perpetua. Esto disminuye el valor de la patente para el propietario —aunque no por mucho, debido al descuento a valor presente,[6] y al hecho de que la mayoría de los productos y procesos se remplazan después de tantos años— y, por lo tanto, el monto de los recursos que se dedicarán a la obtención de patentes. Además, las patentes se obtienen a menudo antes de que se complete el desarrollo comercial (retrasarse en la obtención de la patente es arriesgarse a que otro inventor se adelante), en cuyo caso el periodo de la patente comercialmente valiosa será menor de 20 años. Y el dueño de la patente debe revelar públicamente el invento. Aunque los competidores no pueden copiarlo, sí pueden aprender por la revelación cómo eludir la patente o, en otras palabras, cómo crear un sustituto cercano que no la viole. Por lo tanto, la revelación reduce el costo para los competidores, lo que a su vez ejerce una presión competitiva sobre el dueño de la patente.
2. El hecho de que las patentes se otorguen antes de que un invento haya sido llevado hasta el punto de la viabilidad comercial evita la duplicación de un trabajo de desarrollo costoso.[7] Empero, si se otorgan demasiado temprano —antes de que el inventor sepa efectivamente cómo realizar el producto o proceso que constituye el invento—, sería posible que una patente retardara la innovación al desalentar a otras empresas de desarrollar la tecnología necesaria para comercializar el invento.[8]
3. Los inventos no son patentables si son "obvios". El significado funcional de la obviedad puede descubrirse a bajo costo.[9] Entre menor sea el costo del descubrimiento, menos necesaria será la protección de las patentes para inducir al descubrimiento. Recordemos que la finalidad de la protección de la patente es crear incentivos para inventar al per-

[6] Por ejemplo, con una tasa de descuento de 10%, si se recibe un dólar en 21 años se tiene un valor presente de sólo 14 centavos. Con una tasa de descuento de 5%, el valor presente de un dólar que se recibirá en 21 años es por la cantidad más respetable de 36 centavos, pero probablemente la tasa de descuento más alta es más realista para una actividad arriesgada como la de inventar.

[7] Kitch, nota 4 anterior. Adviértase la analogía con el principio del "buscador comprometido" mencionado en la sección anterior.

[8] Véase Robert P. Merges y Richard R. Nelson, *Market Structure and Technical Advance: The Role of Patent Scope Decisions, in Antitrust, Innovation, and Competitiveness*, 185 (Thomas M. Jorde y David J. Teece [comps.], 1992).

[9] Michael Abramowicz y John F. Duffy, "The Inducement Standard of Patentability", 120 *Yale L. J.*, 1590 (2011); Edmund W. Kitch, "Graham *vs.* John Deere Co.: New Standards for Patents", 1966 *S. Ct. Rev.*, 293; Roberts *vs*. Sears, Roebuck & Co.: 723 F.2d, 1324, 1344 (7th Cir., 1983) (en banc) (opinión dividida).

mitir que el inventor recupere sus costos hundidos de la invención; si esos costos hundidos son ligeros, no hay razón para otorgar una patente:

> Si la innovación se creara y se revelara incluso sin protección de patentes, negar una patente sobre la innovación no tiene costo para la sociedad (porque la innovación se desarrollaría de cualquier modo) y evita que la sociedad sufra innecesariamente las bien conocidas consecuencias negativas de las patentes, incluyendo la restricción a la producción ocasionada por los derechos exclusivos del dueño de la patente y los costos administrativos y de litigio asociados con el funcionamiento de un sistema de patentes […] Una invención debiera considerarse obvia cuando [una persona con conocimientos ordinarios en la materia] considera obvio que valdría la pena intentar los experimentos necesarios para obtener la invención en cuestión, incluso si fuera imposible obtener una patente sobre los resultados o mantenerlos como un secreto comercial.[10]

Desafortunadamente, ni los investigadores de patentes ni los jueces y jurados consideran los costos hundidos al decidir sobre la patentabilidad.[11]

4. El solicitante de una patente debe demostrar no sólo que su invento no es obvio (y que es novedoso: si se ha ideado antes, la concesión de una patente compensaría excesivamente al solicitante; él no habría sido añadido al catálogo de ideas útiles), sino también que es "útil". Podría pensarse que este requerimiento es superfluo. Si el invento no es útil, no generará rentas monopólicas significativas. Pero la eliminación de los inventos inútiles reduce el costo de la búsqueda de patentes para inventores subsecuentes. Sin embargo, es una falacia el deseo de establecer un estándar elevado de utilidad a fin de limitar las rentas monopólicas de las patentes;[12] entre más útil sea el producto, mayores, no menores, serán las rentas.
5. Las ideas fundamentales (por ejemplo, las leyes de la física) no son patentables, a pesar de tener con frecuencia un gran valor. Hasta el advenimiento de los costosos aceleradores de partículas atómicas, la investigación básica casi nunca implicaba gastos sustanciales, de modo que la protección de las patentes podría haber generado una investigación básica excesiva. Sin embargo, la no patentabilidad de los descubrimientos básicos, como la duración limitada de las patentes, refleja algo más que una preocupación por los costos de la adquisición de patentes; hay también severos problemas de identificación, como en el caso de los

[10] Abramowicz y Duffy, nota 9 anterior, en 1594, 1663 (se omiten las notas al pie de página).

[11] Alexander Tabarrok, "Patent Theory *vs.* Patent Law", 1 *Contributions to Econ. Analysis and Policy*, 2002, https://mason.gmu.cdu/-atabarro/PatentPublished.pdf.

[12] El error cometido en Brenner *vs.* Manson, 383 U.S., 519 (1966).

> animales salvajes. Una idea no tiene un lugar físico estable, como una parcela de tierra. Con el paso del tiempo se vuelve cada vez más difícil la identificación de los productos en los que está incorporada una idea particular, en especial si se trata de una idea básica que tiene muchas y variadas aplicaciones. Además, por la definición de "investigación básica", es probable que las oportunidades comerciales se encuentren lejanas en el futuro. Por lo tanto, en términos estrictamente financieros, el valor presente de una patente sobre la investigación básica sería por lo general demasiado pequeño para motivar a una empresa lucrativa a realizar la investigación. Sin embargo, las universidades y otras instituciones de investigación básica buscarían patentes para obtener regalías de otros investigadores.

Estas medidas para evitar las ineficiencias potenciales del sistema de patentes han resultado inadecuadas. El sistema de patentes proporciona a muchos inventores un aislamiento de la competencia mayor de lo necesario para ser incentivado a realizar la invención. El resultado es aumentar los precios de mercado por encima de los niveles eficientes, provocando distorsiones en la asignación de recursos; ocasionar competencias de patentes que conllevan un desperdicio: un desperdicio debido a la duplicación de esfuerzos y porque son innecesarias para inducir la invención (aunque las competencias sí aumentan el ritmo de la invención); aumentar el costo de la búsqueda de los registros de la Oficina de Patentes y Marcas Registradas para asegurarse de que la invención no esté violando la patente de otra persona con su invención; fomentar la presentación de patentes defensivas (debido a la anticipación de que alguien más patentara un producto similar y lo acusara a uno de violación), lo que puede retrasar la innovación; crear gastos de litigio e incertidumbre empresarial y, en ocasiones, la muerte; y alentar a los troles de patentes (discutidos al final de esta sección).

No es que se deba condenar por completo a las patentes. Incluso las competencias de patentes, por más ineficientes que sean, aceleran realmente la invención. Y las patentes publican información valiosa. Se debate cuán valiosa; pero lo que está claro es que cuando la protección de patentes no está disponible las empresas recurren a menudo al secreto comercial (véase la siguiente sección de este capítulo), que, por definición, no implica la publicación y, por lo tanto, limita la difusión de información tecnológica. Además, sin una patente, el inventor podría tener grandes dificultades para obtener el financiamiento que necesita para comercializar su invención; la patente certifica (aunque de forma imperfecta) que la invención es nueva, útil y poco obvia. Se dice que una patente es para un financiero de empresas nuevas lo que una escritura es para un prestamista de hipoteca: prueba de la garantía valiosa.

A fin de evaluar la protección óptima de la patente para un invento específico se requiere considerar el costo de la invención y el costo de la copia (es

decir, el costo de producción y comercialización para una empresa que no sea el inventor o el licenciatario del inventor). Cuanto mayor sea la relación del primero al último, mayor será la protección óptima de la patente para el inventor. La proporción es muy alta para los fármacos. Éstos son, de hecho, la referencia emblemática para la protección de patentes. El costo de inventar un medicamento nuevo, un costo que incluye pruebas extensas que se requieren para la aprobación del medicamento para la venta, está en los cientos de millones de dólares, y, sin embargo, para la mayoría de los medicamentos el costo de copiarlos —o de producir un sustituto idéntico— es muy bajo. Por lo tanto, el inventor podría no ser capaz de recuperar sus costos iniciales sin la protección de patentes (y por las razones adicionales de que el valor presente de los ingresos de la venta del medicamento se ha reducido debido al periodo de tiempo que se tarda en obtener la aprobación, y a que el término real de la patente se trunca debido a que se le confiere concesión y el periodo de protección de la patente comienza a correr cuando se otorga la patente, en lugar de años más tarde, cuando el medicamento pueda comenzar a venderse). Sin embargo, parece haber abusos graves del derecho de patentes en la industria farmacéutica, incluyendo acuerdos colusorios con fabricantes de genéricos (véase después XXI.8) y patentes de "continuación" que amplían la protección de patentes mucho más allá del término de 25 años de una patente farmacéutica.

Hoy en día, el problema de la protección excesiva de las patentes se ilustra mejor por la industria del software. Ésta es una industria progresiva y dinámica pletórica de inventos. Sin embargo, las condiciones que vuelven esencial la protección de patentes en la industria farmacéutica están ausentes. En la actualidad, la mayoría de la innovación de software es incremental, la crean equipos de ingenieros de software a un costo módico y es también efímera —la mayoría de las invenciones de software se remplazan rápidamente—. Además, la innovación de software tiende a darse de forma asistemática. Las patentes no cubren dispositivos completos sino componentes, de modo que un dispositivo de software (un teléfono móvil, una tableta, una computadora portátil, etc.) puede tener decenas o incluso cientos de miles de componentes separados (unos bits de códigos de software o bits de hardware), cada uno patentado. El resultado es una enorme maraña de patentes (se dice que hay al menos 400 000 patentes de software vigentes actualmente, es decir, que no caducaron ni se invalidaron), lo que crea muchas oportunidades para los esfuerzos por incapacitar a los competidores al demandar por violación… y también por violar y, por ende, cuestionar la validez de la patente cuando el dueño de ésta lo demanda a uno.

Otras barreras para una política de patentes eficaz en la industria del software incluyen la escasez de investigadores de patentes en la Oficina de Patentes y Marcas Registradas que tengan las habilidades técnicas necesarias, la dificultad de hurgar los registros de patentes en busca de las patentes de

software pertinentes debido a la falta de un vocabulario técnico estandarizado, la competencia técnica tan limitada de la mayoría de los jueces y jurados, la dificultad para evaluar el pago de daños u otra compensación por la violación de un componente en lugar del producto completo (a menudo es difícil o incluso imposible estimar la contribución del componente al valor que tiene el producto para los consumidores) y la volatilidad de la industria del software debido a su dinamismo tecnológico, lo cual crea incentivos tanto para patentar a la defensiva como para violar patentes y, de este modo, aumenta los costos legales.

Un fenómeno reciente que enfatiza los problemas del sistema de patentes de los Estados Unidos es el aumento de los "troles de patentes". El término (para el cual el sinónimo refinado es "entidad no practicante": ENP) se refiere a empresas que compran patentes con el propósito no de ejercer la invención patentada, sino de obtener derechos de licencia de patentes o el pago de daños por violación. Ahora, los troles presentan más de 60% de todas las demandas de patentes. Como se explicó en un artículo reciente, "existen buenas razones para creer que las ENP litigan sus patentes de manera abrumadora a finales del periodo de éstas […] Muchas ENP no presentan sus propias solicitudes de patente, sino que compran patentes en el mercado secundario (a menudo de empresas fallidas) con el fin de litigar. Desde luego que tarda un tiempo que tales patentes lleguen a las ENP. Además, debido a que las ENP valoran las patentes principalmente por su utilidad en la obtención de regalías y el pago de daños de parte de los productores, por lo general estas dueñas de patentes deben esperar hasta que una industria lucrativa se haya desarrollado para presentar una demanda y continúan presentando demandas mientras queden blancos en la mira con el bolsillo lleno".[13]

Los troles se defienden con el argumento de que proporcionan un mercado para los pequeños inventores, que no pueden solventar una demanda para hacer valer sus patentes. Sin embargo, los dueños de patentes que no pueden producir sus invenciones patentadas o hacer que otros se interesen en producirlas bajo licencia son, en su mayoría, beneficiarios de la emisión indiscriminada de patentes de la oficina encargada. Un inventor que no puede vender su patente de puerta en puerta porque nadie cree que su invención encontrará un mercado no debiera haber obtenido una patente. Si una invención patentada nunca se utiliza para producir algo, ¿por qué debiera importar si la viola alguien que está produciendo algo? El efecto principal del reino del trol pa-

[13] Brian J. Love, "An Empirical Study of Patent Litigation Timing: Could a Patent Term Reduction Decimate Trolls without Harming Innovators?", 161 *U. Pa. L. Rev.,* 1309, 1317 (2013). Por lo tanto, el dominio de los troles podría reducirse mediante el cobro de tarifas altas para "mantener" las patentes, es decir, para evitar que expiren antes del término actual de 20 o (para patentes de medicamentos) 25 años. Véase James Bessen y Brian J. Love, "Make the Patent 'Polluters' Pay: Using Pigovian Fees to Curb Patent Abuse", Boston University School of Law, Law & Econ. Research Paper Series, núm. 1323 (2013).

rece ser aumentar el costo de la invención al aumentar el riesgo para los inventores serios de ser demandados sobre la base de patentes débiles que sólo se usan como armas para obtener derechos de licencia y el pago de daños de parte de ellos. "Las patentes se utilizan para forzar acuerdos irrazonables, ocultar la conducta embarazosa, presionar a otros para que hagan entrega de derechos y participar en esquemas anticompetitivos conocidos como corso. (El corso se produce cuando una empresa productora transfiere sus patentes a una entidad de defensa [es decir, un trol] para que la entidad pueda ir tras los competidores de la compañía.)"[14]

Una forma de reducir la concesión de patentes de bajo valor sería elevar el precio de la solicitud de patente,[15] aunque eso desalentaría algo que valiera la pena inventar; algunos pequeños inventores tendrían dificultades para financiar su actividad inventiva porque sus costos serían mayores.

XI.3. El secreto comercial[16]

Un fabricante que confía en que podrá mantener en secreto su proceso de fabricación durante más tiempo del que podría protegerlo con una patente podría optar por recurrir al secreto y olvidarse de la búsqueda de una patente. Para un secreto comercial no hay ninguna limitación de tiempo, lo que puede parecer raro ya que la persona que tiene el secreto no necesita probar que satisface los criterios del derecho de patentes: novedad, no obviedad y todo lo demás. Sin embargo, la ley del secreto comercial sólo impide la apropiación *ilegal* (por ejemplo, por el cuasidelito o el incumplimiento de contrato) del secreto; los competidores están en libertad de duplicarlo mediante el descubrimiento independiente, o incluso por la ingeniería en reversa del producto del poseedor del secreto, así como de aprovecharse de cualquier revelación accidental del secreto por parte de este último. Si el secreto puede descubrirse fácilmente mediante esfuerzos inventivos independientes, el tenedor del secreto tendrá poco que ganar del gasto de recursos en el mantenimiento de un secreto, mientras que, si éste es tan original e ingenioso que difícilmente podría ser descubierto por el esfuerzo independiente dentro del periodo de protección de la patente, la protección más prolongada que confiera el secreto comercial proveerá una recompensa debidamente incrementada por la creatividad extraordinaria.

[14] Robin Feldman, "FTC Must Move Quickly against Patent Trolls", San Francisco Chronicle, 25 de agosto de 2013, www.sfchronic1e.com/opinion/article/FTC-must-move-quickly-against-patent-trolls-4676067.php.

[15] Como se recomienda en David Fagundes y Jonathan S. Masur, "Costly Intellectual Property", 65 *Vanderbilt L. Rev.,* 677 (2012).

[16] Véase David D. Friedman, William M. Landes y Richard A Posner, "Some Economics of Trade Secret Law", *J. Econ. Perspectives*, invierno de 1991, p. 61.

Si el secreto puede desenmascararse sólo mediante una inversión considerable, pero el gasto sería redituable en vista de los beneficios esperados, el tenedor del secreto tendrá que preocuparse de que si no patenta su proceso o producto secreto lo hará un competidor. Porque, después de la expiración del periodo de un año de gracia para la solicitud de la patente de un invento que su creador ya ha empezado a usar, el inventor no lo podrá patentar por sí mismo (si ha mantenido el invento en secreto) ni impedir que un descubridor independiente lo patente. Por lo tanto, podemos esperar que el tenedor de un secreto gaste recursos considerables en el mantenimiento de su secreto sólo en los casos inusitados en los que es improbable que un competidor, aun a gran costo, descubra el secreto por sí solo; y, si la inutilidad del esfuerzo es obvia, el competidor no incurrirá en los gastos y no habrá desperdicio de recursos.

Cuando la innovación es asequible y muchas personas tienen las habilidades necesarias (¡en la industria de la computación esto incluye a los adolescentes con coeficientes de inteligencia muy altos!), es probable que haya muchas innovaciones simultáneas. Por lo tanto, depender del secreto comercial para proteger la propia innovación será arriesgado: es probable que al menos uno de los innovadores simultáneos busque una patente.[17] Patentar se convierte en una maniobra defensiva en lugar de un estímulo a la innovación.

Como se señaló brevemente en la sección anterior de este capítulo, las patentes tienen una ventaja importante sobre el secreto comercial desde el punto de vista del fomento a la innovación: al publicarse, y contar con detalle suficiente para permitir la reproducción de la invención patentada y crear así "invenciones alternas", la patente fomenta innovaciones adicionales.

XI.4. La privacidad[18]

El secreto no sólo figura en el derecho del secreto comercial sino también en el derecho de la privacidad, que se trata convencionalmente como una rama del derecho de los cuasidelitos, pero existen afinidades con el secreto comercial y, por lo tanto, se puede considerar como una forma de la propiedad intelectual. Tanto el secreto comercial como la privacidad implican el mantenimiento del secreto. La diferencia es que el método o proceso productivo que oculta una empresa crea valor, y el incentivo para crear ese valor se esfumaría si los competidores pudieran descubrir la innovación rápidamente y duplicarla. Sin embargo, un problema característico del derecho de la privaci-

[17] Klaus Kultti, Tuomas Takalo y Juuso Toikka, "Simultaneous Model of Innovation, Secrecy, and Patent Protection", 96 *Am. Econ. Rev. Papers & Proceedings*, 82 (mayo de 2006).

[18] Véanse sobre la economía del derecho a la privacidad Richard A. Posner, *The Economics of Justice*, caps. 9-10 (1981); Richard A. Posner, *Overcoming Law*, cap. 25 (1995); Richard S. Murphy, "Property Rights in Personal Information: An Economic Defense of Privacy", 84 *Geo. L. J.*, 2381 (1996).

dad consiste en saber si una persona debiera tener derecho a ocultar hechos vergonzosos acerca de sí misma. Ya no hay un apoyo judicial considerable para tal derecho. El economista percibe un paralelo con los esfuerzos de los vendedores para ocultar los defectos de sus productos. Un individuo "se vende" a sí mismo tratando de persuadir a sus socios en transacciones potenciales —un patrón, una novia, incluso un conocido casual— de que es honesto, leal, confiable y, en suma, virtuoso. ¿Debiera alentársele para que engañe a estas personas otorgándole un derecho de demandar a quien revele su falta de virtud? Sería diferente si lo que se revelara fuese una receta para algún platillo sensacional. Éste sería un caso de protección como un secreto comercial para alentar la inversión en la creación de una pieza valiosa de propiedad intelectual. La privacidad tiene valor, pero no necesariamente valor social, ya que a menudo se utiliza para engañar.

El primer reconocimiento judicial de un derecho de la privacidad explícito surgió en un caso en el que el demandado había utilizado el nombre y la imagen del demandante en un anuncio sin su consentimiento. Paradójicamente, esta rama del derecho a la privacidad la invocan con mayor frecuencia las celebridades ávidas de publicidad (y, por lo tanto, en ocasiones se la llama "derecho de publicidad"); sólo quieren asegurarse de obtener el mayor precio posible por el uso de su nombre y fotografía. De modo que, en el derecho, la "privacidad" incluye a la publicidad de manera confusa. (El término "derecho a la privacidad" se utiliza, también de forma confusa, para señalar el derecho constitucional a practicarse un aborto.)

Podría parecer que la creación de un derecho de publicidad no conduciría a ninguna inversión que valiera la pena socialmente, sino que tan sólo enriquecería a las celebridades que ya son ricas. Sin embargo, cualquier valor informativo que tenga el respaldo de una celebridad para los consumidores se perdería si cada anunciante pudiera usar el nombre y la imagen de la celebridad. Al igual que en el caso de los pastizales comunales (véase antes III.1), el valor social agregado de relacionar el nombre de la celebridad con un producto en particular puede disminuir si se permite a otros utilizar el nombre en relación con sus productos.

La existencia de una externalidad de congestión proporciona el argumento de que los derechos de publicidad debieran ser perpetuos y, por lo tanto, heredables, lo cual es la tendencia en el derecho.[19] No queremos que esta forma de información o expresión sea del dominio público si ahí será menos valiosa, tanto si la celebridad está viva como muerta. Por ejemplo, si cualquiera fuera libre de incorporar al personaje de dibujos animados "Mickey Mouse" en un libro, película, canción, etc., luego de que expirara el *copyright*, el valor comercial del personaje se desplomaría.

[19] Véase Mark F. Grady, "A Positive Economic Theory of the Right of Publicity", 1 *UCLA Entertainment L. Rev.,* 97, 103, 126 (1994).

XI.5. EL *COPYRIGHT*

El derecho de registro de marca (*copyright*) se asemeja al derecho de las patentes por cuanto otorga derechos limitados en el tiempo, pero se parece al derecho del secreto comercial por cuanto permite el descubrimiento independiente. La razón de esta última característica podría ser que las patentes protegen sólo los inventos, los que pueden ser, y muchos de los cuales (no todos) son ordenados según un índice en la Oficina de Patentes y Marcas Registradas, mientras que los *copyrights* protegen una infinidad de oraciones, frases musicales, detalles de planos arquitectónicos y otras minucias de la expresión, lo que imposibilita por su costo la búsqueda del cuerpo entero de materiales registrados pertinentes a fin de asegurarse de que no se está cometiendo una infracción; por lo tanto, es inevitable cierta cantidad de copia inadvertida.

La limitación temporal de los *copyrights* es ahora tan generosa (la vida del autor más 70 años) que podemos preguntarnos por qué no llega la ley al extremo de otorgar *copyrights* perpetuos. El peligro de atraer recursos excesivos hacia la producción de obras registradas no puede ser la explicación; a resultas del descuento al valor presente (véase antes VI.11), es improbable que el conocimiento de que puede tener derecho a una regalía sobre su libro un siglo después de su publicación afecte su comportamiento el día de hoy. Pero, si los derechos de propiedad sobre la tierra son perpetuos, ¿por qué no sobre los libros? La respuesta es el carácter de "bien público" de la propiedad intelectual, señalado anteriormente. Mientras que, idealmente, *todas* las tierras debieran ser propiedad de alguien, a fin de impedir el surgimiento de las externalidades del congestionamiento que hemos discutido en conexión con los pastos naturales, el uso de alguna pieza de información por parte de A no volverá más costoso el uso de la misma información por parte de B. Ése es un argumento a favor de un gran dominio público de la propiedad intelectual; sin embargo, mientras más largo sea el periodo de *copyright*, menor será el dominio público.

Este punto sugiere no sólo una preocupación por los periodos largos del *copyright*, sino también la necesidad de calificar la demanda de los dueños de la propiedad intelectual de que la violación de los derechos de propiedad intelectual es una forma de "robo". Supongamos que alguien copia de manera ilegal un software protegido por el *copyright*, pero que la copia sólo es para su uso personal y no para la reventa y que no le habría alcanzado para comprar el software; entonces el dueño del *copyright* no pierde ni una sola venta a resultas de la "piratería". Cierta cantidad de piratería puede beneficiar realmente al dueño del *copyright*. Puede crear una demanda de productos complementarios fabricados por él por parte de los piratas; desalentar la formación de un mercado de copias legales o ilegales (porque la piratería es una alternativa a la compra de copias así como de los originales y, por lo tanto, reduce

los mercados de copiadores); y, al acelerar la difusión del software, ayudar a su creador a obtener un monopolio de red (véase antes X.11).

Aun así, hay bases económicas incluso para abogar por los *copyrights* perpetuos. Consideremos a Mickey Mouse de nuevo. Si fuese ubicuo, el público no sólo se cansaría pronto de él, sino que su imagen se volvería borrosa porque los distintos escritores, anunciantes, etc., le darían características diferentes. En la medida en que tales apropiaciones del personaje de Mickey Mouse se clasificaran como parodias, éstas estarían protegidas por la doctrina del uso justo incluso si el *copyright* fuese perpetuo: ello debido a que la duración y el alcance de un *copyright* son cuestiones distintas; y el uso justo, discutido más adelante, es un problema de alcance. Pero no todas estas apropiaciones se clasificarían así, por lo que la doctrina del uso justo no las protegería de la responsabilidad bajo un régimen de *copyright* perpetuo (o indefinidamente renovable, para minimizar los costos del rastreo).

Otra forma de enunciar este punto es que un enfoque exclusivo sobre los incentivos para la creación *inicial* de una obra expresiva puede ser demasiado limitado. Imaginemos una novela, publicada hace muchos años, cuyo *copyright* ha expirado. Se redescubre al novelista y hay un aumento en la demanda de sus novelas. Dado que ningún editor podría establecer un derecho de propiedad sobre las novelas, los incentivos de los editores para publicarlas y promocionarlas podrían ser inadecuados desde un punto de vista social. Además, un editor tendría un incentivo para hacer cambios en la novela, ya que podría registrar el *copyright* de los cambios. Sin embargo, los cambios realizados sólo para reclamar un derecho no son eficientes. A la inversa, si la novela fuera oscura debido a su edad y necesitara un aparato académico elaborado, reedición u otras mejoras costosas para hacerla accesible a un público moderno, los editores podrían estar renuentes a tomar las medidas necesarias, temiendo que no se pudiera recuperar el costo frente a la competencia de las ediciones baratas y básicas de la novela.

Por otro lado, si bien es natural suponer que el alcance de los derechos de propiedad intelectual (incluida su duración) refleja un esfuerzo por encontrar un equilibrio entre los intereses de los creadores y los de los usuarios de la propiedad intelectual, los propios creadores podrían beneficiarse de la imposición de límites a esos derechos. (Esto se relaciona con el punto anterior de que alguna piratería de obras protegidas por *copyrights* puede beneficiar realmente al dueño de tales registros.) La mayoría de los poemas, las novelas, las obras de teatro, las composiciones musicales, las películas y otras obras creativas se basan en gran medida en obras creativas anteriores: se toman prestados algunos detalles de la trama, personajes convencionales, metáforas, progresiones armónicas, ángulos de cámara, etc., de las obras anteriores. Entre mayor sea el alcance de la protección del *copyright* para las obras anteriores, mayor será el costo de la creación de obras subsecuentes. Un aumento del alcance de la protección del *copyright* aumentará los ingresos

esperados por un autor de la venta o el licenciamiento de sus propios *copyrights*, pero también aumentará *su* costo de creación de las obras que él registra. El trueque entre beneficios y sacrificios favorece a una limitación temporal de la protección del *copyright*. Mientras que el aumento del valor presente derivado de un incremento de los ingresos en el futuro distante tenderá a ser insignificante, el aumento en el costo de un autor podría ser grande si, debido al *copyright* perpetuo, ninguna de las obras anteriores fuera del dominio público y, por ende, no estuviera disponible para ser usada en la creación de obras nuevas sin necesidad de incurrir en costos de *copyright*.

Estos costos serían de tres tipos. Los *costos de rastreo* son los que resultan de la identificación del tenedor del *copyright*, quien podría ser el heredero remoto del autor original.[20] Los *costos de negociación* son los que implica la negociación de una licencia con el tenedor del *copyright*. Los *costos de asignación* surgen del hecho de que, en la medida en que la obra registrada sea un bien público puro, cualquier precio positivo de su uso podría inducir a la sustitución por algo cuya producción cueste más a la sociedad o sea de calidad inferior. Esto supone, aunque realistamente, que el tenedor del *copyright* no puede discriminar perfectamente en el precio; es decir, no puede cobrar precios diferentes a compradores diferentes de tal modo que ningún comprador se vaya sin comprar si está dispuesto a pagar un precio por encima del costo de oportunidad (que sería igual a cero en el caso de un bien público que puede copiarse a un costo de cero).

La identificación del tenedor del *copyright* sería fácil y, por lo tanto, los costos del rastreo serían bajos, si la ley requiriera que los dueños de *copyrights* los registraran de nuevo regularmente en algún registro central con el nombre del tenedor del *copyright* que apareciera en la obra registrada. Entonces, una búsqueda del nombre de ese tenedor revelaría quién es el tenedor actual del *copyright* de quien tendría que buscarse una licencia. En cuanto a los costos de negociación, en la mayoría de los casos la obra nueva que infrinja una obra muy antigua (en un régimen de *copyright* perpetuo) infringirá una sola obra (por ejemplo, *Ulises* y la *Odisea*, *Amor sin barreras* y *Romeo y Julieta*, *Ragtime* y *Michael Kohlhaas*), de modo que sólo será necesaria una negociación entre dos partes. Con todo, esto supone que el dueño del *copyright* puede identificarse fácilmente, lo que es difícil con el sistema actual, en el que el *copyright* se reivindica en vez de registrarse. (Los Estados Unidos solían utilizar un sistema de registro, pero se vieron obligados a abandonarlo como condición para tomar parte en el Convenio de Berna, que crea la protección internacional del *copyright*.)

La doctrina del uso justo es un buen ejemplo de cómo la limitación de los derechos del propietario de un *copyright* puede aumentar efectivamente

[20] Véase William F. Patry y Richard A. Posner, "Fair Use and Statutory Reform in the Wake of *Eldred*", 91 *Cal. L. Rev.,* 1639 (2004).

su valor (es otro ejemplo de la función económica de la negación de un derecho de propiedad). Por ejemplo, la doctrina permite que el reseñador de un libro cite pasajes del libro sin obtener el permiso del tenedor del *copyright*. Esto disminuye el costo de las reseñas de libros y, por lo tanto, aumenta su número. Los autores se benefician como grupo, ya que las reseñas de libros son una publicidad gratuita. Aun las reseñas desfavorables estimulan las ventas, por lo menos cuando la alternativa sería la ausencia de toda reseña. Pero la mayoría de las reseñas son favorables, porque a la mayoría de la gente le interesa más que le digan lo que debe leer que lo que no debe leer. Las reseñas de libros son una publicidad particularmente *creíble*, además, porque no están controladas por el anunciante (es decir, el editor del libro). Si los autores pudieran censurar las reseñas de sus libros negando el permiso para citarlos, las reseñas de libros no serían más creíbles que la publicidad pagada. Por lo tanto, los autores sufrirían como grupo si se rechazara el uso justo para las reseñas de libros, aunque algún autor ganara ocasionalmente.

Asimismo, cuando la reseña de un libro disminuye sus ventas no lo hace porque, como una infracción rutinaria del *copyright*, provea a la demanda del libro —raras veces es la reseña de un libro un sustituto cercano del libro reseñado—, sino porque señala las fallas del libro y así provee una información valiosa sin minar las remuneraciones de la creación de una propiedad intelectual *valiosa*. La clase de daño que sufre un autor cuando se llama la atención sobre la *falta* de valor de la propiedad intelectual que ha creado no es la clase de daño que debiera tratar de impedir una ley que trate de alentar la producción de propiedad intelectual.

En el caso Sony, la Corte Suprema sostuvo que la doctrina del uso justo permite la venta de videograbadoras para su uso en la grabación de programas de televisión, aunque no se pague ninguna regalía a los dueños del *copyright* por el privilegio de la grabación.[21] Muchas personas usan sus videograbadoras para grabar programas que se exhiben a una hora inconveniente o que desean ver más de una vez. Tales usos benefician a los propietarios del *copyright*, aunque no se pague ninguna regalía. La mayoría de los programas son comprados por anunciantes que pagan más entre mayor sea el número de los televidentes a los que lleguen; al aumentar la audiencia efectiva de un programa, una videograbadora permite que el dueño del *copyright* cobre más a los anunciantes. Dado que la videograbadora se utilizaba de forma que no constituía una violación sustancial, la Corte Suprema declaró que Sony no era responsable como infractor contributivo (por ejemplo, al contribuir a la violación del *copyright* por parte de personas que utilizaban sus videograbadoras para borrar los comerciales del programa que grabaron).

[21] Sony Corp. of America *vs.* University City Studios, Inc., 464 U.S., 417 (1984).

La Corte llegó a un resultado diferente en el caso Grokster.[22] Grokster distribuyó, de forma gratuita, un software que permitía a los dueños de computadoras formar redes "de pares", por lo que podía transmitirse la copia de un archivo en una computadora a otra. Grokster fue demandada por los dueños de los *copyrights* de música y películas. Señalaron que un propietario del software Grokster podía buscar en la red compuesta por otros propietarios del software un archivo de computadora que contuviera música o películas con *copyright* y, si lo encontraba, copiar tal archivo. Esto constituiría una violación, y la Corte Suprema dictaminó que existía suficiente evidencia de "violación contributiva" de parte de Grokster para ordenar un juicio.

Un infractor contributivo es alguien que facilita la infracción. Al hacerlo responsable, la ley reduce el costo del cumplimiento del *copyright*; sería muy costoso para los dueños del *copyright* demandar a todas las personas involucradas en el intercambio de archivos. El enfoque general no se limita al derecho del *copyright*. Alguien que vende herramientas de ladrón a un ladrón es culpable de ayudar y ser cómplice si el ladrón utiliza las herramientas para cometer un robo. Y alguien que provoca un incumplimiento de contrato es responsable ante la víctima del incumplimiento, aunque la víctima pudo haber demandado a la otra parte del contrato por incumplimiento de contrato. Sin embargo, las dificultades surgen cuando un producto tiene usos tanto legales como ilegales. Sería ridículo responsabilizar a los vendedores de cuchillos de carnicero civil o criminalmente si un comprador usara el cuchillo para cometer un asesinato, a menos que el vendedor supiera que ése era el uso previsto.

Pueden existir usos legales del software de intercambio de archivos. No todas las grabaciones musicales o películas cuentan con *copyright* (éstos podrían haber expirado); y, además, algunos productores de obras registradas, especialmente si son novatos e intentan establecer una reputación (por ejemplo, una banda nueva de rock and roll), pueden desear que sus obras se distribuyan tan rápido y tanto como sea posible incluso sacrificando las regalías. En efecto, estos productores están dispuestos a otorgar licencias libres de regalías a los usuarios del software Grokster. Además, al igual que en el caso de las reseñas de libros, algunas "violaciones" en forma de intercambio de archivos pueden beneficiar efectivamente al tenedor de un *copyright* y, aun así, éste puede negarse a otorgar una licencia libre de regalías porque cree que puede obtener regalías y tener así lo mejor de ambas cosas. La gente pagará más por un CD si cree que puede compartirlo con otros a través de una red de pares, porque entonces el CD se convierte en un tipo de moneda para arreglar intercambios ventajosos e incluso "comprar" amigos nuevos. Entonces, también, si viajas mucho y llevas una computadora contigo, con el intercambio de archivos no tienes que arrastrar tus CD favoritos contigo; en tal situación, el uso compartido de archivos te da simplemente un acceso mejor a un producto por el que

[22] MGM Studios, Inc. *vs.* Grokster, Ltd., 545 U.S., 913 (2005).

pagaste… y pagarás más por el producto si tienes un mejor acceso, al igual que pagarías más por un teléfono celular que pudieses usar en cualquier país del mundo. Además, a través del intercambio de archivos puedes desarrollar un gusto por una banda o un compositor en particular cuyas grabaciones no habrías comprado sin la exposición gratuita a ellos; esto es el intercambio de archivos como publicidad de grabaciones con *copyright* que es gratuita para el tenedor de un *copyright*.

Sin embargo, la Corte citó pruebas que indicaban que la mayor parte del uso compartido de archivos posibilitado por el software de Grokster, como Grokster sabía bien, era violatorio, al igual que la mayoría de las ventas de herramientas de ladrones para ladrones lo son para éstos, aunque algunas son para personas que quieren entrar en su propia casa si olvidan sus llaves y otras son para la policía, o para los curiosos, o para los aspirantes a ladrones que se echan para atrás antes de cometer su primer robo con las herramientas nuevas. La infracción contributiva sería una caja vacía si los usos legales, por escasos que fuesen en relación con los usos ilícitos, permitieran que el infractor contribuyente saliera librado en todos los usos de su producto a pesar de que supiera que en su gran mayoría eran violatorios. Con todo, como alternativa al requisito de que el sistema judicial realizara una evaluación caso por caso respecto a si debe imponer responsabilidad por la infracción contributiva a empresas similares a Grokster, los tenedores de *copyrights* que quisieran poder demandar por infracción contributiva podrían estar obligados a colocar una etiqueta electrónica no removible en sus cd que una computadora hubiera de leer, identificando el cd o archivo descargado a partir de él como contenedor de material protegido. Los productores de software estarían exentos de responsabilidad por infracción contributiva si diseñaran su software para impedir la copia de un archivo etiquetado.

Otro ejemplo de la manera en que podría beneficiarse el tenedor de un *copyright* por una limitación de la protección del *copyright* es el de las representaciones públicas de una obra con *copyright*. Podría parecer obvio que el *copyright* de una canción grabada debiera extenderse a cualquier interpretación de la canción; por ejemplo, por un cantante en una fiesta de bodas.[23] Ahora bien, si los invitados a la fiesta no han oído previamente la canción, es posible que su interpretación los lleve a salir a comprar la grabación de esa canción. Por supuesto, el tenedor del *copyright* estaría en mejor situación si los invitados hicieran tal cosa *y* si el anfitrión de la fiesta pagara una regalía por el derecho de interpretación. Pero, debido a que mucha de la música de las fiestas de bodas es del dominio público, el anfitrión podría no pagar ninguna regalía y, aun así, dejar de interpretar la canción por el conocimiento que tiene de que hay severos castigos para los infractores del *copyright*; entonces el tenedor

[23] Véase Bruce H. Kobayashi y Ben T. Yu, "An Economic Analysis of Performance Rights: Some Implications of the Copyright Act of 1976", 17 *Research in Law and Econ.*, 237 (1995).

del *copyright* estaría en mejor situación si la protección del *copyright* no se extendiera a tales interpretaciones.

Podemos generalizar a partir de estos ejemplos con la ayuda de la distinción que establece el economista entre complementos y sustitutos. Un producto es un complemento de otro si una baja de su precio hace que aumente la cantidad demandada del otro producto, y es un sustituto si una baja de su precio propicia que disminuya la cantidad demandada del otro producto. Un desarmador y un tornillo son complementarios; un desarmador y un martillo son sustitutos. Una reseña favorable de un libro es un complemento del libro reseñado (y hemos visto que no hay manera de confinar el privilegio del uso justo a las reseñas favorables); la videograbadora y la interpretación pública son complementos y sustitutos a la vez. El efecto dominante determinará si los tenedores del *copyright* ganan o pierden con la invocación del privilegio del uso justo a fin de limitar el alcance de la protección del *copyright*.[24]

Dadas las limitaciones del alcance de la protección del *copyright*, ¿crea el derecho del *copyright algún* problema de búsqueda de renta, análogo al creado por el derecho de las patentes? La respuesta es sí, sobre todo en el caso de las extensiones retroactivas del *copyright*, como en la Ley de Extensión del Periodo de Copyright Sonny Bono de 1998, que extendió el *copyright* sobre Mickey Mouse y otros *copyrights* lucrativos que estaban próximos a vencer. Dado que se incurre en los costos principales de creación de la propiedad intelectual antes de la venta de la propiedad, los ingresos de la extensión de un *copyright* son casi una ganancia pura (renta económica).

Pero, al igual que las patentes, los *copyrights* aumentan los costos de transacción; otro productor puede duplicar la obra patentada o registrada sólo si ese productor obtiene una licencia. En al menos un campo, el de las publicaciones académicas distintas de los libros de texto, los costos de transacción parecen ser una pérdida. Incluso si tales obras no estuvieran registradas,[25] los académicos seguirían publicando, no sólo por aquello de "publicar o morir" (los académicos enseñan sólo unas pocas horas a la semana, pero se les paga un salario completo, y al menos la mitad es para compensarlos por llevar a cabo investigación académica y escribir los resultados), sino también porque obtienen ingresos no monetarios significativos por publicar los resultados de la investigación. Y, dado que muy pocas publicaciones académicas son remuneradoras, es probable que la pérdida de los ingresos por *copyrights* tenga un efecto menor en la cantidad de dichas publicaciones.

[24] ¿Dónde cabe la parodia en este análisis? Por un lado, una parodia es como una reseña humorística o crítica de un libro. Por otro lado, puede funcionar como un sustituto de la obra parodiada, por ejemplo, *El joven Frankenstein*, una parodia de la película original de *Frankenstein*. Véase Richard A. Posner, "When Is Parody Fair Use?", 21 *J. Legal Stud.*, 67 (1992).

[25] Como se insiste en Steven Shavell, "Should Copyright of Academic Works Be Abolished?", 2 *J. Legal Analysis*, 301 (2010).

XI.6. El movimiento de código abierto

El movimiento de software de "código abierto" se refiere a la colaboración voluntaria de miles de programadores, quienes no reciben pago ni derecho de propiedad al crear un código de software nuevo. El sistema operativo Linux, el navegador Mozilla Firefox y el servidor Apache son ejemplos exitosos de proyectos de código abierto. Los miembros de los grupos programadores que trabajan en proyectos concretos acuerdan, de hecho, no reivindicar derechos de propiedad intelectual sobre sus contribuciones. Como consecuencia, el código fuente que crean y mejoran está "abierto" para quien quiera usarlo. Suele haber un comité que decide qué contribuciones formarán parte del código fuente "oficial".

¿Por qué razón un programador experimentado dona su tiempo, de forma gratuita, para la creación de software de código abierto cuando podría vender su tiempo a un productor de software comercial como Microsoft? Esto parece desafiar la lógica económica. Sin embargo, la apariencia puede ser engañosa. Puede haber motivaciones económicas para contribuir al software de código abierto.[26] En primer lugar, algunos de los contribuyentes son usuarios del software, y al arreglar errores en él hacen más fácil su propio trabajo como usuarios. En segundo lugar, para la comunidad del software es más fácil evaluar la contribución de un programador al software de código abierto que al código fuente registrado, ya que éste suele mantenerse como secreto comercial. Por lo tanto, puede haber beneficios de carrera por participar, ya sea temporalmente o medio tiempo, en el movimiento de código abierto. En tercer lugar, los programadores expertos obtienen utilidades de su trabajo, y la utilidad puede ser mayor si pueden elegir los proyectos en los que trabajan, como no podrían hacerlo si trabajaran para una empresa comercial. Lo que muestra el movimiento de código abierto no es la incompetencia de la economía para explicar el comportamiento, sino la variedad de incentivos económicos que determinan el comportamiento.

XI.7. Las marcas comerciales

La función económica de las marcas registradas es distinta a la de las patentes y los *copyrights*. Es la de economizar en los costos de búsqueda de los consumidores al asegurar una calidad uniforme. La marca comercial identifica la *fuente* de un producto o servicio particular; por ejemplo, la marca comercial de General Electric identifica a General Electric como el productor

[26] Josh Lerner y Jean Tirole, "Some Simple Economics of Open Source", 50 *J. Industrial Econ.*, 197 (2002).

de los bienes a los que se adhiere la marca comercial. Dado que ésta indica al consumidor a quién culpar si su foco no funciona, otorga a los productores un incentivo para mantener la calidad, lo que a su vez disminuye la necesidad de que el consumidor compre con tanto cuidado como lo haría de otro modo. Aunque el precio nominal de un bien con marca comercial fuese mayor a causa de la inversión del productor en publicidad y mercadeo de su marca, el costo total (los economistas llaman a veces a esto el "precio completo", para distinguirlo del precio nominal, es decir, el precio cobrado por el vendedor) para el consumidor podría ser menor porque la marca comercial transmite información acerca de la calidad cuya obtención podría resultar de otro modo costosa para el consumidor.

El objetivo consiste en capacitar a cada productor para identificar su propia marca sin aumentar los costos de otros productores en la identificación y comercialización de sus marcas. Desde este punto de vista, la mejor marca comercial es la marca fantasiosa, como Kodak: una palabra inventada, por oposición a una palabra tomada del idioma. El número de combinaciones de letras posibles para formar palabras nuevas es prácticamente infinito, de modo que no hay peligro de que una marca fantasiosa aumente el costo de otro productor para encontrar palabras que identifiquen y comercialicen su producto. Mucho más truculenta es una marca "descriptiva"; si se permite a alguien que use "procesador de palabras" como su marca comercial, se volvería costosa para los productores rivales de procesadores de palabras la comercialización de sus marcas porque no podrían usar una descripción compacta. Por lo tanto, la ley protege las marcas descriptivas sólo si la marca ha adquirido el "significado secundario", esto es, sólo si el consumidor ha llegado a identificar el término con una marca particular, y no con el producto en conjunto. Tenemos un ejemplo en "Holliday Inn".

A menudo, cuando un bien de marca comercial inicia su vida como una patente u otro monopolio, la marca registrada llega a denotar al bien mismo, antes que a la fuente, y cuando eso ocurre decimos que la marca comercial es "genérica" y pierde la protección. Algunos ejemplos son los de "aspirina", "celofán" y "yoyo". Si se permitiera que el dueño de la marca comercial excluyera a sus consumidores del uso del término genérico para describir sus marcas, estaría imponiéndoles costos. (Por razones similares, no se permite que un productor registre una característica "operativa" de su producto: El inventor de la rueda no pudo hacer de la rueda su marca y excluir así a la competencia en la fabricación de ruedas.) Sin embargo, una marca comercial no tiene límite temporal ni debe tenerlo. Si hubiera un límite temporal que expirara antes de que el productor hubiese cesado de producir el artículo registrado, tendría que inventar otro nombre para el producto y los consumidores se confundirían. Una marca comercial no puede venderse ni transferirse de otro modo aparte del producto al que designa. Por lo tanto, expira automáticamente cuando el producto deja de fabricarse.

El problema del uso prematuro es analíticamente similar al de la inversión excesiva en el hallazgo de tesoros escondidos (véase antes III.2) o la obtención de una patente. A fin de adquirir un derecho valioso, la gente puede invertir recursos más allá del punto en que tales recursos confieren un beneficio social neto. Construir una casa y cultivar un terreno en tierras públicas es un ejemplo.[27] Si se dispone de la tierra sin un cobro pecuniario, pero el regalo se condiciona (como en efecto ocurrió con la Ley del Homestead de 1862) a la ocupación efectiva de la casa y el cultivo de la tierra, el residente la cultivará hasta el punto en que el último dólar del esfuerzo rinda un dólar de beneficio en la obtención de este derecho, aunque se produzca menos de un dólar de producción agrícola. Por supuesto, si el gobierno *desea* que la tierra sea ocupada a la mayor brevedad posible por razones políticas o militares, una ley que entregue la tierra pública desierta a quien la ocupe y la cultive podría ser un método eficiente para la realización de este propósito.[28]

XI.8. Acuerdos de patentes, las licencias globales de BMI-ASCAP y la intersección entre el antimonopolio y la propiedad intelectual

Los acuerdos de patentes celebrados entre empresas rivales plantean interrogantes difíciles. Tales acuerdos no pueden ser condenados ni aprobados de antemano. Veamos el primer punto. Si dos empresas tienen patentes "bloqueadas", lo que quiere decir que ninguna de las patentes puede emplearse para producir el mejor producto o proceso posible sin infringir a la otra (una consecuencia común de la mejora de patentes, véase antes IX.4), deberá permitirse que las empresas reúnan o compartan licencias de las patentes. Y, si una empresa desarrolla una patente cuyo uso eficiente sólo puede hacerse si se la licencia a empresas rivales (ya que el inventor puede no tener las habilidades o recursos para fabricar el producto o utilizar el proceso que ha inventado), difícilmente podrá prohibírsele a su dueño que la licencie. Supongamos que el titular de una patente es el fabricante de un producto o servicio patentado y no sólo el propietario de la patente, de modo que competirá con sus licenciatarios. En principio, él puede fijar una tasa de regalía que lo proteja contra la competencia triunfante de sus licenciatarios, pero en la práctica podría tener que fijar un precio mínimo al que pudieran vender los licenciatarios a fin de protegerse.

Para explicar esto, supongamos que el costo marginal de A, el dueño de la patente, de producir un adminículo patentado es una suma constante de un

[27] Véase, por ejemplo, Terry L. Anderson y Peter J. Hill, "The Race for Property Rights", 33 *J. Law & Econ.*, 177 (1990). Véase en general David D. Haddock, "First Possession *versus* Optimal Timing: Limiting the Dissipation of Economic Value", 64 *Wash. U. L. Q.*, 775 (1986).

[28] Douglas W. Allen, "Homesteading and Property Rights; or, 'How the West Was Really Won'", 34 *J. Law & Econ.*, 1 (1991).

dólar en una corrida de 100 000 adminículos, pero sería mucho más si tratara de producir más. A cree que B podría producir otros 10 000 adminículos a un costo constante de 95 centavos, que los costos de B serían muy elevados a cualquier cantidad mayor producida y que, a una producción de 110 000, el precio de mercado de los adminículos sería de 1.10 dólares. Por lo tanto, A le cobra una regalía de 15 centavos por adminículo a B, de modo que B perderá dinero si cobra menos de 1.10 por cada adminículo que fabrique. Pero supongamos que resulta que B puede producir 10 000 adminículos adicionales (lo que daría un total de 20 000) a un costo de sólo 75 centavos por adminículo. Si lo hace, la producción total de adminículos será de 120 000, de modo que el precio de mercado bajará (por la Ley de la Demanda), digamos que a 1.01 dólares. Sin embargo, B tendrá un incentivo para producir la cantidad mayor, porque después de pagarle a A la regalía convenida todavía ganará 1 200 en lugar de 1 000 dólares al nivel de producción menor. Con todo, A, en lugar de obtener una ganancia de 11 500 dólares (10 centavos en cada uno de los 100 000 adminículos que produce y 15 centavos en los 10 000 adminículos que según se espera producirá B), sólo ganará 4 000 dólares (un centavo en cada uno de sus 100 000 adminículos y 15 centavos en cada uno de los 20 000 adminículos de B).

Es cierto que si A redujera su producción a 90 000 adminículos a fin de mantener el precio de mercado en 1.10 dólares, su ingreso total derivado de la patente, 12 000 dólares, sería mayor aún que lo esperado cuando otorgó la licencia a B, lo que demuestra que siempre resulta más rentable licenciar la producción a un productor más eficiente que producir uno mismo. Pero es posible que A no advirtiera que B estaba produciendo más de la cantidad esperada de adminículos hasta que el precio empezara a bajar. Además, A podría haber incurrido en costos, basados en una producción esperada de 100 000 adminículos para él, que no podrían haberse evitado por completo cuando la producción disminuyó inesperadamente. Ante tales incertidumbres, si se le prohíbe que establezca un piso (es decir, 1.10 dólares) por debajo del precio de B, A podría optar por no licenciar la patente sino fabricar los adminículos él mismo, aunque sea más eficiente que se permita a B la fabricación de algunos de ellos.

Aunque algunas patentes no podrían desarrollarse de modo óptimo sin la cooperación entre los competidores, cuando una patente es "delgada" (es decir, que podría ser muy probablemente invalidada en los tribunales) su licenciamiento provee una oportunidad única para que las empresas se coludan bajo el disfraz de un licenciamiento legal de la patente. General Electric otorgó en cierta ocasión una licencia a Westinghouse para que produjera focos bajo la patente de GE, a un precio mínimo fijado en el acuerdo de licencia. Bien puede ser que la patente de GE haya sido inválida. General Electric cobraba a Westinghouse la tasa de regalía muy baja de 2%... que, sin embargo, aumentaba a 15% si la participación de Westinghouse en el mercado de los focos aumentaba a más de 15% de ese mercado. Estas condiciones le daban a Westinghouse un desincentivo para expandirse en competencia con General Electric;

si se contentaba con una participación pequeña en el mercado, tendría que pagar sólo una regalía pequeña y compartiría las ganancias monopólicas creadas por una estructura de precios no competitiva.[29]

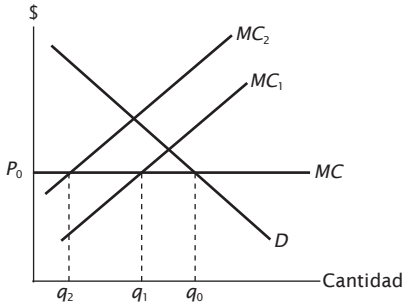
Otro caso de patentes famoso, que pudo haber engañado a la Corte Suprema, es el caso de la "fragmentación", ganado también por los demandados.[30] Varios fabricantes de gasolina habían patentado ciertos procesos para la fabricación de gasolina mediante el método barato de la fragmentación. Estos fabricantes combinaron sus patentes (que aparentemente no eran patentes bloqueadoras) eliminando la competencia entre ellos. La Corte Suprema pensó que éste no era gravemente anticompetitivo porque sólo 26% de la gasolina vendida se producía por el método de la fragmentación y la gasolina producida por otros métodos es físicamente indistinguible, un sustituto perfecto. La gráfica XI.1 muestra que el análisis de la corte estaba incompleto. *CMg* es el costo marginal de producir la gasolina por el método antiguo desplazado (parcialmente) por la fragmentación, y CMg_1 es el costo marginal para los refinadores si los tenedores de las patentes de fragmentación compiten entre sí en el licenciamiento de su uso por los refinadores. Bajo la competencia, la fragmentación es el método más barato hasta la cantidad q_1, más allá de la cual se usará el método comercial para satisfacer el resto de la demanda del mercado (es decir, $q_0 - q_1$). Pero los tenedores de patentes coludidos cobrarán un precio mayor por el uso del método de la fragmentación, lo que eleva los costos marginales de los refinadores al usar ese método hasta algún punto por encima de CMg_1 (en la gráfica XI.1, hasta CMg_2). En consecuencia, se produce menos gasolina por el método de la fragmentación. El precio de la gasolina no cambia porque el método convencional está disponible para producir gasolina adicional a un costo y, por lo tanto, un precio constantes. Pero los costos totales de la producción de gasolina son mayores por la diferencia existente entre las áreas debajo de CMg y CMg_1 entre q_2 y q_1. Esa diferencia es el costo social de la colusión entre los tenedores de las patentes.

Los acuerdos cooperativos entre competidores suelen requerir que el tribunal antimonopolio compense los costos del monopolio con los beneficios de la eficiencia. Consideremos el método[31] por el que las organizaciones de derechos de reproducción en el campo de la música, ASCAP y BMI, comercializan los derechos de reproducción de composiciones musicales con *copyright*. Los compositores (u otros tenedores de *copyrights*) licencian a la organización para que comercialice estas canciones. A su vez, la organización otorga una licencia a las estaciones de radio y otras entidades de reproducción para que usen cualquier canción de su repertorio (cada organización tiene derechos

[29] Sin embargo, la Corte Suprema sostuvo el arreglo. United States *vs*. General Electric Co., 272 U.S., 476 (1926).

[30] Standard Oil Co. (Indiana) *vs*. United States, 283 U.S., 163 (1931).

[31] Que no se consideró ilegal por sí mismo en Broadcast Music, Inc. *vs.* Columbia Broadcasting System, Inc., 441 U.S., 1 (1979).



GRÁFICA XI.1

de comercialización sobre miles de canciones) sin un pago extra. (De ahí la licencia "en blanco"). El único cobro es el de la propia licencia en blanco, un porcentaje especificado de los ingresos del beneficiario de la licencia que no tiene ninguna relación con el mayor o menor grado en que haga uso de sus derechos bajo la licencia. La organización distribuye luego entre los compositores los ingresos obtenidos con sus licencias, aproximadamente en proporción a la frecuencia con la que se reproduzca cada canción.

Cada una de las organizaciones de derechos de reproducción es en efecto una agencia exclusiva de ventas para los compositores que representa y elimina la competencia de precios entre ellos,[32] justo como lo hace la agencia exclusiva de ventas de un cártel convencional. La restricción "en efecto" es necesaria porque el compositor puede, si lo desea, tratar directamente con la estación. ¿Elimina este derecho la posibilidad de precios y ganancias del cártel? En parte. Existen tantos compositores que, si las organizaciones de derechos de reproducción cobraran por la licencia en blanco un precio que excediera el ahorro en costos derivado de dicha licencia, los compositores individuales abandonarían y tratarían directamente con las estaciones, pues cada compositor imaginaría que su engaño no tendría mucho efecto en el mantenimiento del cobro monopólico. Si esta predicción es correcta, el derecho a tratar directamente con la estación pone un tope en el cobro de licencia en blanco equivalente al ahorro en costos de transacción que permite la licencia en blanco. No obstante, esto es un retorno monopólico en cierto sentido, ya que puede exceder los costos de componer música. Aun así, el costo para las estaciones de radio y otros compradores (intermedios) de la música con *copyrights* por tratar individualmente con los compositores sería alto, de modo que el cobro de la licencia en blanco del "cártel" es probablemente menor de lo que sería el cobro correspondiente por los derechos de reproducción musical en un

[32] Sin embargo, un incentivo adicional importante para que un compositor se una a tal organización es que ésta hace cumplir la propiedad intelectual del tenedor del *copyright*. Sería muy difícil (costoso) para un compositor individual detectar violaciones de sus *copyrights* y demandar a los infractores.

mercado puramente "competitivo", lo que ilustra la falacia de igualar a la competencia en el sentido de la asignación eficiente de recursos a los rivales.

Además, la licencia en blanco es un método ingenioso para minimizar los efectos del monopolio sobre la producción. Dado que se permite a cada beneficiario de la licencia reproducir cuanta música quiera sin pagar una cuota adicional, quienes disfrutan de la licencia no tienen ningún incentivo para restringir su uso por debajo del nivel competitivo, como ocurriría con los clientes de un monopolista ordinario.[33] Sin embargo, este método no es perfecto. El cobro de la licencia en blanco podría hacer que algunas estaciones de radio se abstuvieran de reproducir música; incluso podría disminuir el número de estaciones de radio, y se generarían así efectos de sustitución similares a los generados por un cártel convencional.[34]

Otro ejemplo de cómo la política antimonopolio interactúa con la propiedad intelectual se refiere a la decisión de Standard Fashion discutida antes en X.11. Recordemos que Standard Fashion estaba vendiendo patrones de vestidos y, por lo tanto, propiedad intelectual, en lugar de los propios vestidos. Como sabemos, la propiedad intelectual se caracteriza por el tamaño de la diferencia entre el costo promedio y el costo marginal. Supongamos que desarrollar un producto nuevo de software cuesta 10 millones de dólares, el costo marginal es 1 dólar y el precio maximizador de beneficios y la producción anual son 15 y 1 millones, respectivamente. Se espera que el monopolio dure un año si no se impide la entrada, dos años si es así. Si el monopolio dura un año, la ganancia del monopolista será de 4 millones (15 millones – 10 millones (1 millón × 1)). Si dura dos años, su ganancia se eleva a 18 millones (30 millones – 10 millones – (2 millones × 1)). Por lo tanto, debería estar dispuesto a incurrir en costos de hasta 14 millones (de hecho, un poco menos debido a la necesidad de descontar las ganancias futuras a valor presente) para evitar la entrada mediante contratos de exclusividad o métodos de exclusión similares. Por lo tanto, incluso si el gobierno no otorgara monopolios a la propiedad intelectual, como en el derecho de *copyrights* y de patentes, en ocasiones los productores de tal propiedad estarán mejor posicionados para resistir la competencia que los productores de propiedades tangibles, al contrario del supuesto común de que la propiedad intelectual necesita protección monopolística para producirse en absoluto.

Un ejemplo actual de la interfaz antimonopólica de la propiedad intelectual son los acuerdos —el tema de una decisión reciente de la Corte Suprema[35]— relativos a resoluciones en casos de antimonopolio presentados por

[33] En otras palabras, el cobro marginal es cero. Al igual que el costo marginal. No le cuesta más, ni al compositor ni a la compañía de discos, reproducir una vez más una composición grabada. (El deterioro por uso adicional de la grabación o del disco es insignificante.)

[34] Nos encontraremos de nuevo con este problema al considerar la fijación de precios en dos partes después, en XIII.5.

[35] FTC *vs*. Actavis, Inc., 133 S. Ct., 1310 (2013).

productores de productos farmacéuticos de marca comercial (patentados) contra fabricantes de sustitutos genéricos.[36] Una resolución típica dará al fabricante genérico una licencia exclusiva por un periodo de meses. Durante ese periodo, el mercado de ese medicamento específico será un duopolio. El arreglo se asemejará a la licencia de General Electric-Westinghouse discutida en el capítulo anterior.

Lecturas recomendadas

Abramowicz, Michael, y John F. Duffy, "The Inducement Standard of Patentability", 120 *Yale L. J.*, 1590 (2011).

Chien, Colleen, "Predicting Patent Litigation", 90 *Tex. L. Rev.*, 283 (2011).

Devlin, Alan, "Systemic Bias in Patent Law", 61 *DePaul L. Rev.*, 57 (2011).

Economics of Intellectual Property Law (2 vols.; Robert P. Merges [comp.], 2007).

Feldman, Robin, *Rethinking Patent Law* (2012).

Hall, Bronwyn H., y Dietmar Harhoff, "Recent Research on the Economics of Patents", 4 *Ann. Rev. Econ.*, 541 (2012).

Landes, William M., y Richard A Posner, *The Economic Structure of Intellectual Property Law* (2003).

Lemley, Mark A., "The Regulatory Turn in IP", 36 *Harv. J. Law and Pub. Policy*, 109 (2012).

_____, "Distinguishing Lost Profits from Reasonable Royalties", 51 *William and Mary L. Rev.*, 655 (2009).

Mazzoleni, Robert, y Richard R. Nelson, "The Benefits and Costs of Strong Patent Protection: A Contribution to the Current Debate", 27 *Research Policy,* 273 (1998).

Shavell, Steven, "Should Copyright of Academic Works Be Abolished?", 2 *J. Legal Analysis,* 301 (2010).

The Political Economy of Intellectual Property Rights (2 vols.; Christopher May [comp.], 2010).

Problemas

1. La Ley de Inventos de América Leahy-Smith (2011) establece que la patente sobre una invención patentable se concederá no al primer inventor, sino al primer inventor que presente una solicitud de patente. ¿Cuál es el efecto probable del cambio en la ley sobre los inventores en solitario y las grandes compañías, respectivamente? Véase David S. Abrams y R. Polk Wagner, "Poi-

[36] Véanse, por ejemplo, C. Scott Hemphill y Mark A. Lemley, "Earning Exclusivity: Generic Drug Incentives and the Hatch-Waxman Act", 77 *Antitrust L. J.*, 947, 948 (2011); Hemphill, "Paying for Delay: Pharmaceutical Patent Settlement as a Regulatory Design Problem", 81 *N. Y. U. L. Rev.*, 1553 (2006).

soning the Next Apple? The America Invents Act and Individual Inventors", 65 *Stan. L. Rev.,* 517 (2013).

2. Si fuera un autor y recibiera una regalía porcentual fija de su editor, ¿desearía usted que su editor cobrara un precio mayor o menor por el libro en relación con el que el editor desearía cobrar? Relacione su respuesta con las cuestiones de la propiedad dividida que discutimos en el capítulo.

3. Una ley de California otorga a los artistas un derecho (conocido como *droit de suite*) para recibir regalías siempre que su obra de arte se revenda, aun cuando el artista no haya retenido el *copyright* sobre sus obras. ¿Es la ley puramente paternalista? ¿Podrá justificarse económicamente la ley, tal vez, por referencia a beneficios externos o costos externos? Véase Larry S. Karp y Jeffrey M. Perloff, "Legal Requirements That Artists Receive Resale Royalties", 13 *Int'l Rev. Law & Econ.*, 163 (1993).

4. Pueden obtenerse *copyrights* sobre los códigos de computadora. ¿Es esto correcto, aunque el código esté diseñado para ser leído por la computadora antes que por un ser humano?

5. Después de que expira un *copyright*, y la obra registrada vuelve así al dominio público, ¿debiera considerarse a la obra "abandonada" y que un "descubridor" tuviera derecho a registrar un nuevo *copyright* sobre ella?

6. En el caso Cámara de Comercio *vs*. Dow Jones & Co., 456 N.E.2d, 84 (Ill., 1983), la Cámara de Comercio de Chicago creó un contrato de futuros basado en el índice de Dow Jones de 30 acciones industriales. El contrato permitió tanto la especulación sobre los movimientos en este índice ampliamente observado como la cobertura. La Corte Suprema de Illinois sostuvo que la Cámara de Comercio había malversado la inversión de Dow Jones en su índice. Un año después otro tribunal, en el caso U.S. Golf Association *vs*. St. Andrews Systems, Data-Max, Inc., 749 F.2d, 1028 (3rd Cir., 1984), obtuvo el resultado contrario. La USGA, el órgano rector del golf amateur, había desarrollado una fórmula para calcular los hándicaps de los golfistas; los hándicaps permiten que los jugadores con distintas habilidades compitan entre sí sobre una base nominalmente igual. El demandado, Data-Max, obtuvo la fórmula (supuestamente por medios legales) y ofreció un servicio mediante el cual un golfista, por medio de una calculadora portátil, podría obtener un hándicap inmediato y actualizado calculado según la fórmula de USGA. El tribunal se negó a emitir un mandato en contra de Data-Max. Determinó que la apropiación indebida, de ser procesable, requiere de un daño al demandante en su mercado primario, y no sólo en el mercado de un trabajo secundario o derivado. Otra forma de plantear esto es que la apropiación indebida debe amenazar la base de ingresos que sostiene al producto del cual ha sacado provecho gratuito el demandado.

¿Qué decisión tiene más sentido económico, la de Dow Jones o la de Data-Max?

XII. LA REGULACIÓN DE LA RELACIÓN DEL EMPLEO[1]

XII.1. El tratamiento especial de los cárteles laborales

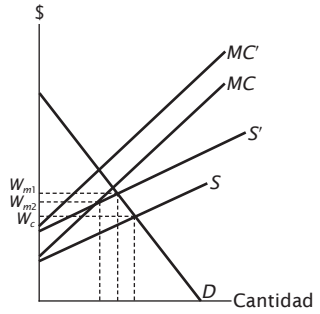
En el siglo XIX, la interrogante principal de la política antimonopólica consistía en saber si debía reprimirse a los sindicatos como combinaciones ilegales que restringían el comercio. Los economistas clásicos pensaban que no,[2] pero tampoco creían que las combinaciones de trabajadores pudieran distinguirse de las combinaciones de empleadores para bajar los salarios o las de vendedores para elevar los precios.[3] El propósito principal de un sindicato, según han creído desde hace largo tiempo la mayoría de los economistas, es el de limitar la oferta de mano de obra para que el empleador no pueda utilizar la competencia entre los trabajadores para controlar el precio de la mano de obra (los salarios). Por lo tanto, el derecho común tenía sólidas razones económicas cuando se negó a hacer cumplir los acuerdos para unirse en sindicatos, prohibió los bloqueos —un intento de interferir con las relaciones contractuales existentes entre la empresa bloqueada y sus clientes, trabajadores u otros proveedores— e hizo cumplir los contratos de perro amarillo (por los que los trabajadores aceptan no unirse a sindicatos durante el tiempo que dura su empleo). Por lo que toca a los últimos, el trabajador demandaría presumiblemente una compensación por renunciar a su derecho de unirse a un sindicato… y, si no era compensado generosamente, ésta no era una pérdida social, ya que toda compensación por no combinarse con otros trabajadores para crear un monopolio laboral es en sí misma una forma de la ganancia monopólica.

Sin embargo, es posible que esta presentación de un derecho común eficiente de las relaciones laborales descanse en supuestos poco realistas acerca de los mercados laborales estadunidenses del siglo XIX. Los empleadores

[1] Por lo que toca a los principios del derecho laboral, véanse Robert A. Gorman, *Basic Text on Labor Law: Unionization and Collective Bargaining* (2ª ed., 2004); *The Developing Labor Law: The Board, the Courts, and the National Labor Relations Act* (Patrick Hardin *et al*. [comps.], 4ª ed., 2002) (2 vols.); Charles A. Sullivan, Rebecca H. White y Michael J. Zimmer, *Cases and Materials on Employment Discrimination* (6A ed., 2003).

[2] Véase, por ejemplo, John Stuart Mill, *Principles of Political Economy,* libro 5, cap. 10, 5 (1848) [edición en español: *Principios de economía política, con algunas de sus aplicaciones a la filosofía social*, trad. de Teodoro Ortiz, rev. de la trad. de C. Lara Beautell, 2ª ed., FCE, México, 2006]. Sin embargo, la opinión general de Mill sobre los sindicatos era muy crítica. Véase Samuel Hollander, "The Economics of John Stuart Mill", vol. 2: *Political Economy*, 897-907 (1985).

[3] A. V. Dicey, *Lectures on the Relation between Law and Public Opinion in England during the Nineteenth Century,* 190-201 (2ª ed., 1914).



GRÁFICA XII.1

tendrían poder monopsónico si los trabajadores ignoraran sus oportunidades alternativas de empleo o tuvieran costos de reubicación muy elevados, o si los empleadores conspiraran para reducir los salarios.[4] Es posible que las tres condiciones anteriores hayan sido comunes en el siglo XIX, cuando había muchos trabajadores inmigrantes, el nivel de la educación era mucho más bajo que ahora, la movilidad de la mano de obra era menor también, las empresas tenían menos competencia porque el costo de transportación era mayor y las leyes antimonopólicas no se hacían cumplir contra las conspiraciones para reducir los salarios. Contra todo esto se encuentran los hechos (algunos de ellos señalados en el capítulo VIII) de que la gran era de la inmigración a este país después de la Guerra Civil fue una respuesta a una demanda de mano de obra en gran expansión, los salarios eran siempre mayores en los Estados Unidos que en otros países y los estadunidenses han sido siempre muy trashumantes. Además, aunque los monopsonios de trabajadores fuesen un problema, los monopolios laborales no son una solución atractiva. La situación es una de monopolio bilateral. Con ambas partes tratando de limitar la oferta de mano de obra, aunque por razones diferentes, la oferta no llegará al nivel competitivo, aunque los salarios serán mayores que si sólo hubiera un monopsonio.

Esto se muestra en la gráfica XII.1 (una amplificación de la gráfica X.3). S es el precio de oferta de la mano de obra antes de la formación de un sindicato, y s_c es el precio competitivo de la mano de obra. El sindicato aparece y eleva el precio de oferta a S' (se supone un precio uniformemente mayor para cualquier cantidad de mano de obra que se provea). Sin monopsonio, la cantidad de mano de obra comprada bajará a q_{m1} y el salario aumentará a

[4] Adam Smith pensaba que tales conspiraciones eran comunes. Véase *The Wealth of Nations* (Edwin Cannan [comp.], 1937 [1776]) [edición en español: *Investigación sobre la naturaleza y causas de la riqueza de las naciones*, ed. de Edwin Cannan, trad. y estudio prelim. de Gabriel Franco, FCE, México, 1958]. Y durante la huelga de Pullman de 1894 el gobierno invocó la Ley Sherman contra los organizadores pero no contra la conspiración de los empleadores para no aumentar los salarios. Almont Lindsey, *The Pullman Strike*, caps. 6-7, 14 (1942).

s_{m1}. Sin embargo, si el empleador es un monopsonista la oferta bajará más aún, a q_{m2}, y el salario bajará a s_{m2}. La solución eficiente implicaría: *1)* que la compañía pagara el salario competitivo y *2)* que hubiera un pago de transferencia hecho de una sola vez, de la compañía al sindicato o del sindicato a la compañía. (¿Qué determinará quién le paga a quién?: las respectivas elasticidades de la demanda y oferta de mano de obra a la empresa. Entre más elástica es la demanda, más fácilmente podrá la firma traer otros insumos como sustitutos y así reducir el salario. Entre más elástica sea la oferta, mayor será la caída en la oferta de mano de obra a la empresa si la empresa intenta reducir el salario.) Pero tales tratos son raros. Un sindicato es un colectivo y, debido a los altos costos de transacción, es difícil repartir tanto la factura de un gran pago al contado como un gran gasto único entre los miembros.

Si dejamos de lado el monopsonio laboral, podemos decir con cierta confianza (pero más adelante, en XII.3, consideraremos una sugerencia en contrario) que el efecto de la sindicalización es la reducción de la oferta de mano de obra en el sector sindicalizado. Los salarios mayores obtenidos por el sindicato inducirán al empleador a sustituir la mano de obra por capital y también mano de obra más cara por otra más barata (por ejemplo, reubicando su empresa en una región del país donde los sindicatos sean débiles) y a trabajadores de cuello azul por trabajadores de cuello blanco. Así pues, algunos trabajadores se benefician con la sindicalización: aquellos a quienes se les pagan salarios mayores en las industrias sindicalizadas y aquellos que son contratados por primera vez por empleadores que buscan sustitutos de la mano de obra sindicalizada. Lo mismo ocurre con algunos accionistas: los de empresas cuyos competidores pagaban antes salarios menores que ellos, pero se ven obligados a pagar el mismo salario a resultas de la sindicalización. También se benefician los consumidores de los productos de empresas no sindicalizadas, porque ahora la oferta de mano de obra para el sector no sindicalizado es mayor, de modo que los salarios debieran ser menores y, por lo tanto, los precios también. Los perdedores son los consumidores de las industrias sindicalizadas (porque los empleadores de esas industrias trasladarán a sus consumidores por lo menos una parte de sus costos laborales mayores; véase antes X.10), los accionistas y los proveedores de esas industrias, los trabajadores que pierden sus empleos debido a la reducción de la demanda de mano de obra causada por las escalas salariales del sindicato,[5] los trabajadores del sector no sindicalizado porque ahora es mayor la oferta de trabajadores en ese sector y los consumidores en general porque los insumos de mano de obra se están empleando menos eficientemente por toda la economía.

Pero supongamos que el sector sindicalizado es muy intensivo en capital. Cuando los salarios aumentan en ese sector, los precios aumentarán y la

[5] Los trabajadores de las minorías podrían verse afectados más adversamente en este sentido. Véanse después XII.7 y XII.9.

producción bajará. Habrá cierta sustitución de capital por mano de obra, pero si el efecto de producción domina al efecto de sustitución podría haber un flujo neto de capital del sector sindicalizado al no sindicalizado. Si es así, aumentará la productividad de la mano de obra en el último sector. La inversión de capital eleva el producto marginal de los trabajadores que operan los bienes de capital. Estos bienes permiten a más trabajadores generar la misma producción o más, de modo que la producción por trabajador es mayor. El aumento salarial debido a este aumento de la productividad podría concebiblemente dominar a la disminución salarial debida al aumento de la oferta de mano de obra; si es así, el resultado conducirá a un aumento neto de los salarios en el sector no sindicalizado. Pero tanto el capital como la mano de obra se emplearían menos eficientemente que si no se hubiese sindicalizado ninguna parte de la economía.

Como podrían haber pronosticado los economistas clásicos, la Ley Sherman se aplicó a las actividades de los sindicatos, en particular en la huelga de Pullman de 1894.[6] Aun después de que se promulgó la Ley Clayton en 1914, la cual eximió al trabajo de las leyes antimonopólicas, algunos tribunales estatales continuaron prohibiendo las huelgas. Pero en la década de 1930 la política pública dio un brusco viraje. La Ley Norris-LaGuardia, de 1932, eliminó virtualmente la prohibición anterior y la Ley Wagner de 1935 contenía fuertes protecciones legales a la formación de sindicatos, impulsando un aumento dramático de la actividad sindical y de las tasas salariales en diversas industrias.[7] Aunque la política prosindicalista de la Ley Wagner fue recortada por las enmiendas Taft-Hartley de 1947, la Ley Nacional de Relaciones Laborales (como se conoce a la Ley Wagner enmendada por Taft-Hartley) continúa incorporando una política de aliento a los sindicatos.

Sin embargo, como vimos en el capítulo x, los monopolios y los cárteles llevan en sí mismos las semillas de su destrucción. Los premios salariales de los sindicatos (incluidos los beneficios de franja negociados por el sindicato y las protecciones laborales) elevan los costos marginales de las empresas sindicalizadas, lo que las lleva a perder negocios a manos de las empresas no sindicalizadas y genera finalmente, como ha ocurrido en los Estados Unidos y otras economías avanzadas en decenios recientes, una marcada declinación del porcentaje de la fuerza de trabajo que está sindicalizada.[8] En la actualidad, en

[6] Lindsey, nota 4 anterior, caps. 8, 12. Véase también Anderson *vs.* Shipowners Association of the Pacific Coast, 272 U.S., 359 (1926).

[7] H. Gregg Lewis, *Unionism and Relative Wages in the United States: An Empirical Inquiry* (1963).

[8] Peter D. Linneman, Michael L. Wachter y William H. Carter, "Evaluating the Evidence on Union Employment and Wages", 44 *Industrial & Labor Relations Rev.*, 34 (1990); Linneman y Wachter, "Rising Union Premiums and the Declining Boundaries among Noncompeting Groups", 76 *Am. Econ. Rev. Papers & Proceedings*, 103 (mayo de 1986); David G. Blanchflower y Richard B. Freeman, "Unionism in the United States and Other Advanced OECD Countries", 31 *Industrial Relations,* 56 (invierno de 1992).

los Estados Unidos menos de 7% de los trabajadores privados están sindicalizados, en comparación con 35% en 1950, aunque más de un tercio de los trabajadores públicos están sindicalizados en comparación con 12% de entonces. El porcentaje de todos los trabajadores que hoy están sindicalizados es 11%, un tercio de lo que era en 1950 y poco más de la mitad de lo que fue en 1980.

La disparidad actual entre la sindicalización de los empleados públicos y los privados refleja el hecho de que los empleados públicos son votantes y, por lo tanto, sus empleadores son reacios a contrariarlos a través de la lucha contra los esfuerzos de sindicalización; los empleadores públicos están menos motivados para minimizar los costos, ya que no compiten en el mercado; y, en parte debido a la falta de presión competitiva, el nepotismo, el amiguismo y el favoritismo político son abundantes en el empleo público y, por ende, los empleados creen, a menudo con razón, que necesitan un sindicato que los proteja contra las prácticas arbitrarias del empleo. Por lo que los sindicatos de empleados públicos son, hasta cierto punto, una consecuencia de los incentivos y las limitaciones enquistadas en el empleo público, pero también desempeñan un papel causal en los salarios elevados y los generosos beneficios que caracterizan al empleo público en comparación con el privado.

Desde 1950, la disminución de la sindicalización del sector privado refleja una variedad de factores: Las regulaciones gubernamentales considerables de la seguridad y la salud en los lugares de trabajo han reducido el valor de los sindicatos para los trabajadores, al igual que lo han hecho la movilidad laboral y el carácter cada vez más técnico e individualizado de muchos empleos, lo que dificulta que los trabajadores se pongan de acuerdo en los términos y condiciones de empleo que debieran buscar de manera colectiva. La competencia internacional ha erosionado el poder de los sindicatos para obtener salarios, beneficios y normas de trabajo supercompetitivos, al igual que lo ha hecho el movimiento de desregulación, que ha vuelto a las industrias reguladas anteriormente, como la del transporte, más competitivas. Los sindicatos tienen poco poder en las industrias competitivas, porque el salario supracompetitivo, al aumentar los costos de los empleadores, desplazará su producción hacia los competidores. Ésta es la lección del debilitamiento de los fabricantes de automóviles de Detroit.

El aumento en la sindicalización del sector público puede reflejar en gran medida el aumento del empleo en dicho sector, lo que ha hecho que los empleados de éste tengan mayor influencia política.

XII.2. La lógica económica de la Ley Nacional de Relaciones Laborales

La Ley Nacional de Relaciones Laborales (LNRL) es una especie de Ley Sherman en reversa, diseñada para alentar la cartelización de los mercados laborales. Estudiando cómo hace esto la LNRL, profundizaremos nuestro entendi-

miento de la economía de los cárteles y también veremos que la lógica económica de la ley no es siempre una lógica de eficiencia.

Si la política pública fuese neutral en lo tocante a la sindicalización, los esfuerzos para sindicalizar una planta u otro establecimiento encontrarían serios problemas de gorroneo. Un trabajador que ayudara al sindicato podría ser despedido, ya que, aunque la compañía incurriría en un costo al despedirlo, suponiendo que era un trabajador satisfactorio (pero si no lo era ya hubiese sido despedido), el costo podría ser menor que el beneficio de señalar a los empleados restantes que si cualquiera de ellos tomara el lugar del empleado despedido como organizador del sindicato también sería despedido. Si los trabajadores se apoyaran mutuamente y se fueran a huelga en apoyo del empleado despedido, se alteraría la balanza de los costos y el empleador podría tener que dar marcha atrás. Sin embargo, los trabajadores enfrentarían problemas serios de gorroneo al ponerse de acuerdo sobre la huelga (y acerca de su duración, etcétera).

Pero, como ocurre con los precios depredadores (véase antes X.9), la lógica económica de la lucha entre el empleador y los trabajadores es más compleja que nuestras simples intuiciones. Si el trabajador tiene buenas oportunidades alternativas de empleo, es posible que lo tenga sin cuidado la amenaza de su despido si ayuda al sindicato. Si no tiene buenas alternativas, es probable que ello ocurra porque tiene habilidades especializadas para trabajar para su empleador específico ("capital humano específico de la empresa"; véase después XII.4). Esto lo volverá más productivo de lo que sería si trabajara para otra empresa, lo que implica que estará recibiendo un sueldo mayor del que recibiría si perdiera su empleo y fuera a trabajar en otra empresa. Al mismo tiempo, el costo de despedirlo será mayor para el empleador de lo que sería si el empleado no tuviera un capital humano específico de la empresa.

El sindicato resuelve los problemas de gorroneo que encontrarían los trabajadores si trataran de sindicalizar la planta por sí solos. Pero la solución no es completa, porque los trabajadores se verán tentados a aprovecharse gratuitamente de los esfuerzos del sindicato. Y, si una campaña de organización sindical no puede triunfar sin los esfuerzos enérgicos de algunos trabajadores de la planta, el empleador podría ser capaz de frustrar la campaña de organización a un costo bajo mediante despidos selectivos —aunque la Ley Nacional de Relaciones Laborales prohíbe al empleador despedir o discriminar de otro modo a los organizadores sindicales o simpatizantes—. (Y la Ley Norris-LaGuardia prohíbe los contratos de perro amarillo o antisindicales, que permitirían al empleador explotar los incentivos de los trabajadores para el gorroneo. Cada trabajador tendría un incentivo para firmar tal contrato por una remuneración moderada, sabiendo que, si la mayoría de los trabajadores firmaran, quienes no lo hicieran serían demasiado pocos para ser capaces de sindicalizarse efectivamente (véase antes X.10).)

Si la campaña de organización triunfa hasta el punto de que al menos 30% de los trabajadores firman tarjetas que autorizan al sindicato a ser su representante en la negociación colectiva, la Junta Nacional de Relaciones (NLRB) realizará una elección para representante en la negociación colectiva. Si el sindicato recibe la mayoría de los votos emitidos, se convertirá en el representante exclusivo de los trabajadores para la negociación. El empleador deberá entonces negociar con el sindicato, de buena fe, las condiciones de un contrato laboral para todos los trabajadores de la unidad de negociación colectiva; no se le permitirá negociar por separado con trabajadores individuales.

La elección de representantes, el principio de la representación exclusiva y la afiliación sindical obligatoria[9] constituyen un conjunto ingenioso de instrumentos (que opera un poco como la unificación de los campos de petróleo y gas, discutida en el capítulo III) para superar los problemas del gorroneo que, de otro modo, azotarían como una plaga al sindicato como un cártel de muchos miembros. Si no existieran estos instrumentos, la estrategia racional para el trabajador individual sería abstenerse de unirse al sindicato o de destinar cualesquiera recursos a su promoción. Porque, si el sindicato triunfara gracias a los esfuerzos de otros trabajadores, él tendría el beneficio del salario sindical, pero le iría mejor que a los miembros del sindicato al no tener que pagar cuotas y podría trabajar más horas, si lo desea, a un salario ligeramente menor que el sindical, justo como ocurre con la empresa de la franja en un mercado cartelizado (véase antes X.1). Tal gorroneo es impedido por los principios de la regla de la mayoría y la representación exclusiva; el trabajador individual no puede esperar un tratamiento especial ni negociar condiciones diferentes con el empleador. Otra práctica por la que la ley impide el gorroneo, esta vez por el empleador en relación con los esfuerzos del sindicato, consiste en prohibir que el empleador ofrezca a los trabajadores, durante la campaña electoral, salarios mayores o mejores condiciones de trabajo.

El diseño de la unidad electoral en la elección de representante, lo que se llama la unidad negociadora, es decisivo. La Junta de Trabajo certificará a cualquier grupo de empleados que sea a la vez homogéneo en lo tocante a las condiciones del empleo (salarios, prestaciones accesorias, deberes laborales, etc.) y distinto de otros empleados de la empresa. Una sola planta o instalación puede contener varias unidades de negociación diferentes, cada una de las cuales regateará por separado con el empleador. De acuerdo con la política

[9] La afiliación sindical obligatoria —que las partes de una negociación contractual colectiva pueden exigir, fuera de los estados donde se aplica el derecho al trabajo— significa que todo trabajador en la unidad de negociación colectiva deberá unirse al sindicato o por lo menos pagar una parte de las cuotas sindicales, lo que se llama un pago de agencia. Se ha descubierto que las leyes del derecho al trabajo tienen efectos negativos sustanciales sobre la sindicalización y los salarios. Gasper A. Garofalo y Devinder M. Malhotra, "An Integrated Model of the Economic Effects of Right-to-Work Laws", 13 *J. Lab Res.*, 293 (1992).

legal de promover los cárteles de trabajadores, la Junta certifica generalmente a la unidad más pequeña posible, antes que a la más grande. Los costos de transacción entre los trabajadores son menores entre menor sea el número de ellos y más armoniosos sean sus intereses. Mientras más trabajadores haya en la unidad de negociación, más difícil será para ellos desarrollar una posición común y mutuamente satisfactoria —será como determinar el precio óptimo del cártel cuando los cartelistas tienen costos diferentes—. Y los beneficios de la sindicalización son mayores entre más pequeña sea la unidad en relación con el total de empleados en la empresa. Supongamos que la unidad consta de 10 trabajadores en una planta de 100, y que todos los trabajadores ganan lo mismo. Entonces, un aumento de 10% para los miembros de la unidad se traducirá en un aumento de sólo 1% en el total de la nómina salarial del empleador. Por supuesto, la unidad debe ser suficientemente importante para el trabajo de la empresa a fin de poder hacer una amenaza de huelga creíble, pero si la unidad realmente tiene un perfil intransferible esta condición podrá satisfacerse, aunque la unidad sea pequeña.[10]

La clave para entender la economía de las huelgas es el monopolio bilateral del empleador y el sindicato. Un cártel de productores disminuirá la producción pero no hasta cero, por supuesto. No obstante, si los clientes se reunieran en un cártel de compradores, o si sólo hubiera un cliente, éste o aquél podrían responder al cártel de los vendedores amenazando con dejar de comprarle, con la esperanza de que la amenaza de tal boicot indujera al cártel a dar marcha atrás. Un empleador es, en efecto, el único comprador de los servicios de trabajo controlados por el sindicato. Si el sindicato anuncia que desea un precio mayor por tales servicios el empleador puede negarse, amenazando en efecto con no comprar nada, y entonces el sindicato, si quiere mantener la credibilidad, deberá decir que el empleador está intimidando con amenazas que no puede o no le conviene cumplir, llamando a los trabajadores a la huelga. Sin embargo, una parte de la obligación legal de negociar de buena fe consiste en abstenerse de anunciar desde el inicio de las negociaciones que la oferta es "firme", algo que la otra parte debe tomar o dejar. Tal anuncio es una medida de compromiso previo. Si ambas partes recurren a esta medida, no podrán contemporizar y estallará una huelga.[11]

Una huelga impondrá costos a ambas partes: salarios perdidos para los trabajadores y ganancias perdidas para el empleador (a menos que haya una gran oferta de trabajadores de remplazo). Por lo tanto, se podría pensar que las partes estarían siempre en mejor situación negociando un arreglo. El problema, ya conocido por nuestra discusión de los precios depredadores en

[10] Douglas L. Leslie ofrece una visión contraria en "Labor Bargaining Units", 70 *Va. L. Rev.*, 353 (1984).

[11] Véase Douglas G. Baird, Robert H. Gertner y Randal C. Picker, *Game Theory and the Law,* 43-45 (1994).

el capítulo x, es que los términos del acuerdo dependen de la credibilidad de las amenazas respectivas de las partes. La disposición del empleador a aceptar una huelga ocasional, y la del sindicato a llamar a una huelga ocasional, podrían generar beneficios de largo plazo (*ex ante*) para ambas partes por cuanto al aumento de la credibilidad que superarían a los costos de la huelga a corto plazo. Debido a que "es muy fácil hablar", "acompañar las palabras con acciones" es la forma más creíble de señalización. Esto supone que las huelgas son más probables a principios de la relación negociadora o, de sobrevenir más tarde, cuando hay un cambio dramático en las circunstancias. En ambos casos, las partes tenderán a carecer de buena información sobre el precio de reserva de cada una y será difícil para ellas obtener buena información sin que recíprocamente una rete a la otra a que haga cumplir sus bravuconadas. Será difícil para el empleador saber si la amenaza de huelga del sindicato es vana a menos que éste se niegue a ceder y, de igual manera, será difícil para el sindicato saber si la amenaza del empleador de aceptar una larga huelga es vana a menos que el sindicato llame a huelga.

Una huelga impondrá costos al empleador sólo si no puede remplazar a los trabajadores huelguistas de inmediato. La Ley Nacional de Relaciones Laborales facilita el remplazo de los huelguistas en tres formas, y lo dificulta en tres formas. Lo facilita negando la protección de la ley frente a los empleados de supervisión, permitiendo así que el empleador los utilice como remplazos temporales; permitiendo que el empleador contrate remplazos permanentes (la contrapartida de los miembros recién ingresados a los mercados de productos cartelizados); y prohibiendo que los huelguistas dañen la propiedad del empleador (por ejemplo, abandonando el puesto sin apagar la maquinaria). Dado que al dañar la propiedad los trabajadores podrían imponer costos muy grandes al empleador con costos mínimos para ellos (en comparación con los grandes costos de una huelga prolongada para los trabajadores), si se les permitiera que dañaran la propiedad aumentaría grandemente el poder monopólico de los sindicatos.

Sin embargo, la ley dificulta que el empleador opere con trabajadores de remplazo prohibiéndole pagar un salario mayor que el de los trabajadores huelguistas a quienes remplazan, permitiendo que los trabajadores organicen piquetes en la planta y prohibiéndole que rescinda la relación de empleo con los trabajadores huelguistas. El empleador deberá, pues, reinstalar a los huelguistas cuando termine la huelga a menos que los puestos de todos ellos hayan sido llenados con remplazos permanentes, pero en ese caso deberá colocar a los huelguistas a la cabeza de la cola para llenar las vacantes a medida que surjan. Estas tres reglas operan conjuntamente permitiendo que los huelguistas identifiquen a los trabajadores de remplazo, recordando a estos últimos que cuando termine la huelga podrán encontrarse trabajando al lado de los huelguistas —una proximidad incómoda, a veces peligrosa, que disuade a muchas personas de contratarse como remplazos— e impidiendo que el

empleador pague un salario superior para compensar a los remplazos por este costo adicional del trabajo.

XII.3 La Ley de la Libre Elección del Empleado

En 2007, la Cámara de Representantes aprobó la Ley de la Libre Elección del Empleado, una ley que promueve el sindicalismo. El proyecto de ley fracasó en el Senado debido a la oposición de los republicanos. En 2009 se reintrodujo, pero ninguna de las cámaras lo aprobó. Al escribir esto, ya es historia, pero sigue siendo interesante analíticamente.

La disposición más controvertida de la ley habría prescindido del requisito de una elección secreta para determinar si el empleador debía reconocer a un sindicato como representante de los trabajadores. El empleador tendría que reconocer al sindicato, si el sindicato obtenía las tarjetas de autorización firmadas por la mayoría de los trabajadores en la unidad de negociación. Otra disposición controvertida exigía que si dentro de un plazo determinado (incluido un periodo de mediación) el sindicato y el empleador no pudieran llegar a convenir sobre los términos de los acuerdos de negociación colectiva, su disputa se sometería a arbitraje obligatorio. Por consiguiente, los árbitros determinarían esos términos.

La disposición de la firma de tarjetas habría facilitado mucho más para los sindicatos la organización de empresas en las que existiese apoyo sindical considerable pero no el suficiente para asegurar la victoria en una elección secreta. Los trabajadores que apoyan la sindicalización tienden a sentirse más fuertes que los trabajadores que se oponen a ella y, así, tendrían más posibilidades de ejercer presión sobre los indecisos para que firmaran las tarjetas que las concedidas a los oponentes para instar a los dubitativos a que no las firmaran. En algunos lugares de trabajo la presión incluiría amenazas de violencia.

El arbitraje obligatorio también promovería la sindicalización. Eliminaría los costos de la huelga en contra de un empleador terco y atraería a los trabajadores porque se podría esperar que un árbitro fuera más generoso que el empleador al establecer los términos y condiciones de trabajo, ya que no apoyaría ningún aumento en los costos laborales, aunque habría casos en los que el sindicato podría obtener más concesiones de un empleador yendo a huelga o amenazando con hacerlo que las que un árbitro probablemente le concedería.

La elección del momento en que se impulsó la Ley de la Libre Elección del Empleado fue poco favorable. En 2009 el país se estaba hundiendo en la depresión. Lo último que necesita un país en una depresión es un aumento del sindicalismo. Los sindicatos aumentan el desempleo al obtener aumentos salariales que reducen la producción de los empleadores aumentando los costos. Y la respuesta de los empleadores ante un aumento en los costos laborales es el despido de trabajadores, en lugar de una reducción de salarios;

los recortes salariales desmoralizan a toda la fuerza de trabajo, lo que conduce a una menor productividad, mientras que los despidos afectan a menos trabajadores y podrían estimular a los trabajadores restantes a trabajar más para evitar ser despedidos.

Los sindicatos reducen la flexibilidad de gestión, lo que es particularmente perjudicial en una depresión o recesión. Los contratos sindicales suelen requerir que los despidos sean estrictamente inversos a la antigüedad y, como resultado, el empleador no puede utilizar la necesidad generada por la depresión para despedir trabajadores y deshacerse de gente inútil; por esa razón y debido a la inflexibilidad salarial creada por los contratos sindicales, las empresas sindicalizadas no pueden reducir sus costos en respuesta a una caída en la demanda de sus productos tan rápida o profundamente como lo pueden hacer las empresas no sindicalizadas, y esto retrasa el restablecimiento del equilibrio económico. El efecto es particularmente notable en una deflación, lo que es un grave peligro en una depresión o una recesión severa. Una deflación es una reducción en la cantidad de dinero en circulación; puede ser resultado de la acumulación de dinero de la gente debido al miedo de correr cualquier riesgo financiero. La reducción en la cantidad de dinero en circulación disminuye la relación de dinero a bienes y servicios, de manera que los precios caen (la misma cantidad de dinero compra más). De ahí que los salarios nominales constantes aumenten los salarios reales (el mismo salario monetario compra más), incrementando así el costo de los productores y provocando más despidos. Este efecto fue muy marcado en la Gran Depresión de 1930.

El incoherente programa del New Deal de lucha contra la depresión combinó medidas sensatas como salirse del patrón oro, ampliar la oferta monetaria y aumentar el empleo mediante la adopción de programas de obras públicas con programas de restricción de la producción como la Ley Nacional de Recuperación Industrial, que fomentó la formación de cárteles de productores, y la Ley de Ajuste Agrícola, que reducía la producción agrícola para aumentar los ingresos de los agricultores… y la Ley Wagner, que fomentaba la formación sindical.

XII.4. Los sindicatos y la productividad

Algunos economistas han especulado que la sindicalización podría aumentar la productividad, contrarrestando sus efectos monopólicos.[12] Los contratos de negociación colectiva establecen generalmente un mecanismo de quejas para arbitrar las reclamaciones de los trabajadores que también les otorga seguridad en su empleo; no se trata de una seguridad absoluta, porque pueden ser despedidos si disminuye la demanda de mano de obra de la empresa, pero sí una seguridad contra el despido sin causa justificada (determinada por medio

[12] Véase, por ejemplo, Richard B. Freeman y James L. Medoff, *What Do Unions Do?* (1984).

del mecanismo de quejas). Se afirma que, sin un mecanismo de quejas, los empleadores no podrían descubrir que sus capataces estaban maltratando a los trabajadores antes de que ellos advirtieran que la rotación de trabajadores era anormal, y los trabajadores más viejos no compartirían su experiencia con los más jóvenes, por temor a que los más jóvenes los remplazaran. En otras palabras, un sindicato puede reducir los costos de mediación (costos derivados de los conflictos de intereses frecuentes entre, de un lado, los empleados y otros agentes y, del otro, sus empleadores u otros directivos) monitoreando el comportamiento de los supervisores sobre los cuales el empleador tiene un control limitado porque no vigila su trabajo. Otras formas en las cuales los sindicatos pueden incrementar la productividad serían comunicando las preocupaciones de los trabajadores a la administración, ayudando a los trabajadores a adquirir las habilidades necesarias para su progreso, proveyéndolos de beneficios sociales y, en síntesis, mediando entre la mano de obra y la gerencia.

La teoría no explica por qué, si el mecanismo de quejas y la seguridad en el empleo más otras actividades sindicales reducen la costosa rotación y aumentan la eficiencia del trabajador, los empleadores no adoptan estos instrumentos sin esperar a que un sindicato aparezca en escena. (Desde luego, algunos lo hacen.) Si sólo un empleador de una industria cayera en la cuenta de sus ventajas, la competencia obligaría a los demás a seguirlo. Es posible, con todo, que se necesite la intervención de un tercero responsable del cumplimiento para que el mecanismo de quejas o las disposiciones de seguridad en el empleo sean creíbles para los trabajadores. Pero aun si el único tercero creíble fuesen los sindicatos, los empleadores permitirían la sindicalización si ésta aumentara la productividad de su fuerza de trabajo, siempre que la competencia entre los sindicatos redujera el precio de sus servicios al nivel de su costo marginal.

La prueba más contundente en contra de la teoría de que la sindicalización promueve la productividad es que no puede explicar la pronunciada disminución en el porcentaje de trabajadores sindicalizados. Si la sindicalización aumenta la productividad, ese porcentaje debiera estar creciendo.13 Sin embargo, una versión más sutil de la teoría distingue entre dos tipos de sindicalismo, conflictivo y no conflictivo. En el sindicalismo no conflictivo, el sindicato reconoce que está en asociación con el empleador y se centra en actividades que apoyan en lugar de debilitar el funcionamiento eficiente de la empresa. Pero, si el sindicalismo no conflictivo fuera importante, esperaríamos que el sindicalismo hubiera mantenido una mayor vitalidad en el sector privado de la que tiene. De hecho, la mayoría de los sindicatos son conflictivos. Presionan por mecanismos de quejas y seguridad laboral no para promover la producti-

[13] Para más evidencia de que la sindicalización no incrementa la productividad, véase, por ejemplo, Richard S. Ruback y Martin B. Zimmerman, "Unionization and Profitability: Evidence from the Capital Market", 92 *J. Pol. Econ.* 1134 (1984). Para evidencia de que la sindicalización reduce la productividad, véase Barry T. Hirsch, *Labor Unions and the Economic Performance of Firms* (1991).

vidad sino para proteger y favorecer a los partidarios del sindicato. Las reglas de antigüedad negociadas por el sindicato impiden dos cosas: primero, que los empleadores castiguen a los simpatizantes del sindicato despidiéndolos antes que a nadie y, segundo, que den un trato preferencial a un tipo de trabajadores, los de más edad, que son los más aptos para ser los más fuertes partidarios del sindicato, sin importar si son los mejores trabajadores. Los trabajadores más viejos tienen costos de reubicación mayores que los trabajadores más jóvenes, debido a sus lazos más fuertes con el área local construidos durante muchos años y al hecho de que sus habilidades pueden haberse vuelto especializadas para el empleador particular (el capital humano específico de la empresa). Esto otorga al empleador algo de poder monopsónico sobre ellos y, por lo tanto, vuelve al sindicato más valioso para ellos. Asimismo, es más probable que ellos obtengan los beneficios completos de la sindicalización antes que los trabajadores más jóvenes, más móviles, quienes podrían renunciar poco tiempo después de la firma del contrato de negociación colectiva. Y, aunque los trabajadores más viejos tienen un periodo más breve para recibir una ganancia por su apoyo al sindicato, el sindicato puede ajustarse a esto negociando generosos beneficios médicos y de jubilación, lo que por supuesto beneficia principalmente a los trabajadores de más edad. Es cierto que un trabajador especializado con un empleador particular poseerá un poder monopólico además de afrontar un poder monopsónico; la situación es una de monopolio bilateral. Pero su poder es propenso a disminuir con la edad. La edad a la vez reduce la productividad (en el caso de muchos de los trabajadores, aunque no de todos ellos) y el periodo restante en el que el empleador puede esperar beneficiarse de las habilidades especializadas del trabajador.

XII.5. Empleo a voluntad

Otra prueba de que la seguridad del empleo es ineficiente es que, fuera del pequeño y disminuido sector sindicalizado y del empleo gubernamental (donde la antigüedad es común aun en ausencia de sindicatos), el empleo a voluntad es la forma habitual del contrato laboral en los Estados Unidos. El trabajador puede renunciar cuando lo desee, el empleador puede despedir al empleado cuando lo desee. Pudiera parecer que esto dejaría al empleado totalmente a merced del empleador, pero eso no es cierto, porque un empleador que se gana la reputación de despedir arbitrariamente a los empleados tendrá que pagar una prima a los empleados nuevos.

En este ámbito está operando una fuerza más sutil. Los economistas distinguen dos tipos de capital humano[14] (capacidad de ganar ingresos). Uno es

[14] Véase Gary S. Becker, *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education* (3ª ed., 1993). "Capital humano" no es sólo una metáfora. Los economistas emplean el término en el sentido, estrictamente análogo al capital físico, de un activo que

el capital humano general; el otro, con el que ya nos hemos encontrado, es el capital humano específico de una empresa. Los trabajadores que desarrollan habilidades especializadas para un empleador particular son empleados más productivos para este empleador de lo que serían para cualquier otro. La creación de tal capital, como la de cualquier otro capital, es costosa: requiere capacitación en el empleo, además de aprendizaje a través de la experiencia. ¿Quién deberá pagar por esa capacitación? El empleador que paga armará al empleado para que amenace con renunciar a menos que se le pague más, ya que si renuncia se destruirá la inversión que ha hecho el empleador en el capital humano del empleado específico de la empresa. Pero el empleado que paga por la capacitación (comúnmente aceptando un salario menor) armará al empleador para que amenace con despedirlo si no acepta un salario menor, lo que destruirá la inversión del empleado. El incentivo para formular estas amenazas se minimiza si el empleador y el empleado contribuyen al costo de la creación de capital humano específico de la empresa. Porque entonces cada quien tiene menos que perder si el otro termina el empleo y, por lo tanto, es menos probable que se rinda bajo la amenaza destinada a extraer la mayor parte posible del valor creado por la capacitación. Y, si el empleado recibe un salario menor de lo que sería si no estuviera pagando una parte del costo de su capacitación, el empleador perderá dinero al despedirlo, ya que ello terminará con el pago. Esto le da al empleado cierta seguridad en su empleo,[15] al mismo tiempo que economiza en los costos de la redacción y el cumplimiento de un contrato de empleo explícito.[16]

A pesar de su eficiencia, el empleo a voluntad se encuentra bajo el fuego de los tribunales. Un cuasidelito de terminación injusta del derecho común, sensatamente aplicado a los casos en que un trabajador es despedido por ejercer un derecho legal —por ejemplo, prestando un testimonio veraz pero perjudicial contra su empleador en un juicio del gobierno contra el emplea-

genera ganancias a lo largo del tiempo antes que inmediatamente. Sin embargo, las ganancias no tienen por qué ser pecuniarias.

[15] Aunque difícilmente completa. Si la empresa está a punto de cerrar, ni el deseo de preservar su reputación de empleador "justo" ni el deseo de obtener un rendimiento de su inversión en la creación del capital humano de sus trabajadores específico de la empresa le impedirá despedir trabajadores. Pero, por supuesto, si está a punto de cerrar, sus días están contados de todos modos. Adviértase que esta discusión implica que una empresa sucesora debiera estar obligada por un acuerdo de negociación colectiva celebrado por su predecesora; una empresa que está abandonando el mercado tendrá poco o ningún capital de reputación en juego que la pueda obligar a cumplir compromisos implícitos pero no formales.

[16] El capital humano general (dominio de la lengua o alfabetización, conocimiento de matemáticas y de principios básicos de ingeniería, actitudes laborales apropiadas, etc.) tiende a crearse por la escolaridad formal pagada por el individuo, sus padres, el Estado y donantes caritativos, antes que por los empleadores. El capital humano general es portátil, por lo que el empleador que lo otorga no tiene garantía de que cosechará los frutos de una mayor productividad del trabajador instruido. Esto es semejante al capital de infraestructura física (por ejemplo, carreteras que no sean de peaje).

dor por evasión de impuestos—, se está convirtiendo en algunos estados en un requerimiento *de facto* para que se demuestre una buena razón para despedir a un trabajador, aunque sea un empleado a voluntad. Resulta difícil entender cómo pueden beneficiarse de tal requerimiento los trabajadores en general. Si el requerimiento fuera óptimo se negociaría voluntariamente. Ésta no es la clase de problemas de información que podrían frustrar las transacciones acerca de la seguridad en el lugar de trabajo (véase antes VIII.3 y después XII.7), porque los despidos no son acontecimientos tan raros que no pueda esperarse que los trabajadores evalúen racionalmente el riesgo de tener que abandonar el trabajo. El costo para el empleador de demostrar una buena razón para deshacerse de un empleado incompetente debe ser mayor al beneficio para el trabajador de verse asegurado contra un despido injusto por encima del seguro proveído por el capital humano específico de la empresa. Desde luego, el costo extra para el empleador sería un costo laboral y, por lo tanto, disminuiría la cantidad que el empleador estaba dispuesto a pagar en salarios. Pero la respuesta de un empleador a un solicitante de empleo que se ofreciera a aceptar un salario menor a cambio de un contrato de trabajo con duración determinada en lugar de un empleo a voluntad podría ser creer que el trabajador debe ser pesimista acerca de su capacidad para cumplir con las expectativas del empleador.

Es notable el rechazo del empleo a voluntad en la mayoría de los empleos del gobierno. Varios factores están en acción. Si la sociedad, por cuestión de estabilidad política, no quiere que sus servidores públicos corran riesgos, puede atraer a una mano de obra con aversión al riesgo al ofrecer seguridad en el empleo y, al hacerlo, reducir incidentalmente los gastos financieros del servicio público. (En contraste, un empleador privado no quiere atraer trabajadores con aversión al riesgo. Ellos exigirán una prima salarial si se les niega seguridad en el empleo; y un empleador privado es reacio a dar a los trabajadores tal seguridad porque le dificulta adaptarse a los cambios en la demanda de su producto y, por lo tanto, a los cambios en el número y tipo de empleados que necesita.) También porque el público tiende a sospechar de la compensación generosa para los empleados de gobierno, a las agencias gubernamentales les agrada sustituir el salario por la compensación indirecta, como la seguridad en el empleo y generosos beneficios de salud, porque el público puede monitorear los salarios, pero le es difícil asignar un valor en dólares a la compensación no monetaria. Por supuesto, esto implica —pero de forma realista— que los funcionarios del gobierno no son los agentes perfectos del electorado.

Aunque el empleo a voluntad parezca el epítome para la economía del mercado libre, coexiste sorprendentemente con la rigidez salarial. Incluso fuera del sector sindicalizado, los empleadores parecen reticentes a reducir los salarios en respuesta a la caída de la demanda u otra adversidad que reduzca los ingresos de la empresa. Se muestran reacios incluso cuando, en una depresión, el fracaso en la reducción del salario monetario da lugar a un aumento

en el salario real porque el poder adquisitivo de las unidades monetarias está aumentando (eso es lo que significa una deflación). La rigidez de los salarios en sentido descendente es lamentable desde un punto de vista económico global porque aumenta el desempleo y, por lo tanto, reduce la producción y puede causar dificultades económicas graves. (Los subsidios por desempleo mitigan las dificultades, pero también prolongan el desempleo y exacerban los déficits públicos.) Pero la rigidez es racional desde el punto de vista de los empleadores y de la mayoría de los empleados. Es difícil (y probablemente imposible) convencer a los trabajadores de que es necesaria una reducción en su salario monetario para evitar un aumento de su salario real; la diferencia entre lo nominal y lo real es demasiado sutil. Cualquier reducción en su salario les molestará y hará que se preocupen por su situación personal, distrayéndolos de su trabajo. En cambio, despedir a los trabajadores molesta principalmente a los trabajadores despedidos… y se han ido, y su pena no permea el lugar de trabajo. Los trabajadores que no son despedidos están motivados a trabajar más, para no ser despedidos después. Por supuesto, puede ser que no haya suficiente trabajo para todos cuando la demanda de la producción del empleador caiga.

XII.6. La falla del mercado laboral: el caso de los secretarios legales judiciales

He aquí un ejemplo, que debiera ser de gran interés para jueces y estudiantes de derecho, de la manera en que un mercado laboral puede no generar resultados eficientes.[17]

Hace 40 años, los secretarios legales judiciales federales eran contratados en el otoño de su tercer año de la escuela de derecho, menos de un año antes de que iniciaran su empleo en la secretaría. A través de los años, la contratación ha ocurrido más temprano cada vez, y ahora la mayoría de los secretarios son contratados en el otoño de su segundo año de la escuela de derecho, o sea, casi dos años antes del inicio de su empleo en la secretaría. ¿Por qué ha ocurrido esto, dada la clara pérdida de información a causa de que los jueces tomen la decisión de contratación tan temprano, y el potencial resultante para una mala relación entre jueces y secretarios?

La teoría de juegos sugiere la respuesta. El juez intercambia la pérdida de información por la ganancia de ser capaz de contratar un mejor secretario legal que en el caso del que podría conseguir si esperara. El juez obtiene esta ganancia como sigue: ofrece una secretaría a un solicitante prometedor antes de que otros jueces hayan empezado a entrevistar o hacer ofertas, y otor-

[17] Véanse Christopher Avery *et al.*, "The Market for Federal Judicial Law Clerks", 68 *U. Chi. L. Rev.*, 793 (2001); Avery *et al.*, "The New Market for Federal Judicial Law Clerks", 74 *U. Chi. L. Rev.*, 447 (2007). Véase en general Hao Li y Sherwin Rosen, "Unraveling in Matching Markets", 88 *Am. Econ. Rev.*, 371 (1998).

ga al solicitante un tiempo limitado para responder (es decir, hace una "oferta explosiva"). Al decidir si aceptará la oferta, el solicitante intercambia la pérdida de recibir posiblemente ofertas de jueces mejores por el riesgo de no recibir ninguna oferta mejor. Entre más aversión por el riesgo sienta el solicitante, más probable es que acepte la oferta explosiva.

Una vez que un juez adopta esta estrategia y logra obtener mejores secretarios legales, otros jueces se sienten presionados para obrar del mismo modo. El resultado es una carrera para contratar secretarios legales tan pronto como sea posible (aunque en algún momento el juez tendrá tan poca información sobre el solicitante que tendrá que retrasar el momento de la selección; es por eso que el juez no contrataría a un secretario legal antes de que el empleado hubiese comenzado la escuela de derecho, incluso si su expediente universitario y la puntuación en el examen de admisión a la escuela de derecho pronosticaran que sería un buen estudiante. La carrera causa una pérdida de información valiosa, una disminución de las opciones de los solicitantes y un potencial de mala relación entre jueces y secretarios.[18] La mala relación supone que una asignación de secretarios a jueces al azar sería ineficiente, pero probablemente lo sería porque en pocos mercados laborales observamos una asignación aleatoria de los empleados a los empleadores.

Entonces, ¿cuál procedimiento es más eficiente para asignar los secretarios a los jueces: que los mejores jueces tengan a los mejores secretarios o que los peores jueces tengan a los mejores secretarios? El primero es más eficiente si la productividad de jueces y secretarios, considerados como insumos de la generación del producto judicial, es multiplicativa antes que aditiva, como en el ejemplo discutido antes, en v.3. Supongamos que el juez A tiene una productividad de 4 y el juez B una de 2, y que el secretario C tiene una productividad de 3 y el secretario D una de 1.5. Entonces, si se asigna C a A, la productividad total de ambos será de 12; D será asignado a B, y su productividad total en conjunto será de 3. El gran total será de 15. Si por el contrario se asignara D a A y C a B, el gran total sería de sólo 12 ($4 \times 1.5 + 2 \times 3$). Tendría sentido que se asignaran los mejores secretarios a los peores jueces en lugar de a los mejores sólo si los secretarios y los jueces se sustituyeran entre sí en lugar de complementarse. Si se sustituyen, entonces en efecto un buen secretario remplaza a un juez, y queremos que él o ella remplacen a un mal juez y no a uno bueno.

En la primavera de 2002 la mayoría de los jueces federales de apelación estuvieron de acuerdo en no contratar a ningún secretario judicial para 2004 hasta el verano de 2003, es decir, hasta que empezara el último año del estudiante en la escuela de derecho. Hubo un buen cumplimiento del acuerdo el

[18] Se benefician los solicitantes que sientan aversión por el riesgo; el riesgo que se evita es que sus calificaciones de segundo y tercer año no sean tan altas como las de su primer año. Los jueces que sienten aversión por el riesgo se benefician también; el riesgo que evitan es el de no obtener su primera elección como secretarios a menos que se anticipen a la fecha fijada para la contratación.

primer año en que estuvo en vigor, pero el cumplimiento ha disminuido desde entonces. Esto es como lo predeciría la teoría del cártel, porque el acuerdo de los jueces es un acuerdo para no competir y, por lo tanto, un cártel. Al principio los jueces no saben cuántos cumplirán y, puesto que el cumplimiento es de su mayor interés, cumplen, ya que el costo de no cumplir es innegable. Con un buen cumplimiento el primer año, el incentivo para no cumplir aumenta, ya que la mejor posición en la que un juez puede estar es, como en el caso correspondiente a un cártel estándar de vendedores, el de ser el desertor solitario. Por lo tanto, el cumplimiento disminuye en el segundo año, y el proceso continúa en los años posteriores hasta que el cártel colapsa por completo… como sucedió en 2013.

XII.7. Los trabajadores y el derecho antimonopólico

Aunque la cartelización sindical de los trabajadores está exenta de la ley antimonopolio, subsiste un papel limitado para el derecho antimonopólico en el campo laboral: el de impedir que los empleadores utilicen a los sindicatos como agentes en la cartelización de los mercados de productos. Un ejemplo extremo, y obviamente ilegal, sería un acuerdo de negociación colectiva entre un sindicato y una unidad de negociación de varios empleadores que represente a los vendedores principales de la industria para fijar los precios a los que los vendedores deban vender sus productos. Aunque los precios monopólicos de los productos de la industria se traducirían en una reducción de la producción de la industria y, por lo tanto, en una reducción de su demanda de insumos, incluida la mano de obra, las ganancias monopólicas generadas por los precios monopólicos pudieran utilizarse para aumentar los salarios de los trabajadores restantes por encima del nivel anterior, quizá con un sobrante suficiente para compensar a los trabajadores excedentes por los costos de su despido.

A veces no está claro si una disposición restrictiva de un acuerdo laboral trata de aumentar el bienestar del trabajador directamente o de crear un monopolio en los mercados del producto del empleador. Supongamos que un sindicato representante de los trabajadores de la construcción de plomería celebra acuerdos con los subcontratistas de la plomería para no instalar ninguna unidad de aire acondicionado cuya tubería interna haya sido cortada y moldeada en la fábrica antes que por los propios empleados de los subcontratistas. Se ha sostenido que tal acuerdo debe considerarse como creador de un cártel entre los subcontratistas ya que les permite "obtener ganancias mayores si el trabajo demandado al sindicato es realizado por sus compañías antes que por una fábrica".[19] Sin embargo, este argumento es cuestionable, a menos que la prefabricación permita el desplazamiento completo del sub-

[19] Douglas Leslie, "Right to Control: A Study in Secondary Boycotts and Labor Antitrust", 89 *Harv. L. Rev.,* 904, 909 (1976).

contratista. El subcontratista provee un servicio de instalación utilizando dos insumos: mano de obra y equipo de aire acondicionado. Los combinará en las proporciones que minimicen sus costos totales. Si la tubería interna puede ser cortada a un costo menor en la fábrica que por sus propios trabajadores, sus costos de instalación totales se minimizarán comprando los acondicionadores de aire prefabricados. Su incentivo no cambiará si es miembro de un cártel de subcontratistas ya que, como sabemos, las ganancias de un monopolista (o cártel) aumentarán si sus costos bajan. Por lo tanto, los subcontratistas no tienen nada que ganar de un acuerdo para no comprar los acondicionadores de aire prefabricados, y esto implica que su acuerdo para no instalar acondicionadores de aire prefabricados es un instrumento para aumentar la demanda de mano de obra por parte de los subcontratistas, un objetivo válido de la negociación colectiva. Aunque pudiera pensarse que el subcontratista podría compensar al sindicato por el efecto de desempleo del acondicionador de aire prefabricado, ello dependería del monto del ahorro de costos posibilitado por la unidad prefabricada. Supongamos que tal unidad permite al subcontratista prescindir de dos trabajadores con un ahorro de 10 000 dólares por empleo, pero cuesta 9 500 dólares más que el acondicionador de aire fabricado en el sitio de la construcción. El ahorro de 500 dólares podría no ser suficiente para compensar a los trabajadores por los costos de encontrar empleos igualmente buenos en otra parte.

XII.8. La legislación del salario mínimo y de la "protección del trabajador"

La Ley Federal del Salario Mínimo refuerza el efecto de la sindicalización sobre las tasas salariales al limitar la competencia de la mano de obra no sindicalizada. Supongamos que en una parte del país los sindicatos son débiles, de modo que las tasas salariales son menores que en otras partes. Las empresas ubicadas en esa área tendrán una ventaja competitiva que les permitirá crecer a expensas de las empresas de las áreas sindicalizadas, y el resultado será una demanda menor de mano de obra en tales áreas. Una ley de salario mínimo contrarresta esta tendencia al elevar obligatoriamente los salarios de los trabajadores no sindicalizados y reducir así la demanda de esa mano de obra, como se indica en la gráfica XII.2, donde s es el salario competitivo para los trabajadores no sindicalizados y s' es el salario mínimo.

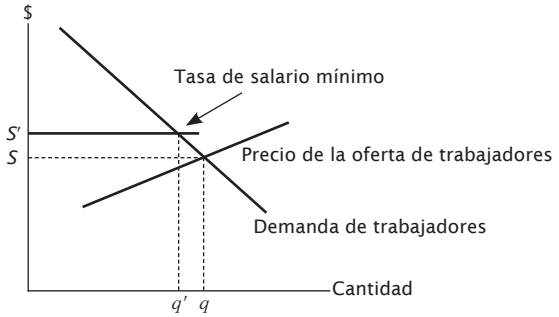
Los efectos de desempleo del salario mínimo se concentran en los trabajadores marginales: mujeres de edad mediana, hombres jóvenes y negros, siendo los más afectados los adolescentes negros.[20] La fijación de un salario

[20] Hay una vasta literatura empírica sobre los efectos económicos del salario mínimo. Véanse, por ejemplo, David Neumark y William Wascher, "State-Level Estimates of Minimum Wage Effects: New Evidence and Interpretations from Disequilibrium Methods", 37 *J. Human Resources*, 35 (2002); Jere R. Behrman, Robin C. Sickles y Paul Taubman, "The Impact of Mini-

mínimo no tiene ningún efecto sobre la demanda de trabajadores cuya productividad marginal sea elevada y que reciben, por ende, un salario de mercado libre por encima del salario mínimo; los desempleados son aquellos trabajadores cuya productividad marginal es menor que el salario mínimo. Dicho de otro modo, el salario mínimo, estando por debajo del salario de los trabajadores sindicalizados, no eleva *sus* salarios por encima del nivel prevaleciente y, por lo tanto, no crea desempleo entre ellos. El salario mínimo aumenta el precio relativo y, por ende, disminuye la atracción de una fuerza de trabajo sustituta constituida por trabajadores marginales que, si no existiera el salario mínimo, compensarían a los empleadores que los contraten aceptando un salario menor que el de los trabajadores sindicalizados más productivos, pero también más caros. Y adviértase que un salario mínimo podría dificultar que los trabajadores marginales obtuvieran capacitación en el trabajo al impedirles compensar al empleador por la capacitación aceptando un salario menor.

Algunos trabajadores, además de los competidores sindicalizados, se ven ayudados por el salario mínimo: aquellos que ofrecen q' horas de trabajo en la gráfica XII.2. Éstos son trabajadores cuyo salario de mercado libre está por debajo del salario mínimo, pero cuyo producto marginal es mayor que dicho salario. (No es que estuvieran mal pagados antes del mínimo decretado, como tampoco es mal pagado un inventor que no logra captar el excedente del consumidor de su invención mediante una discriminación perfecta del

mum Wages on the Distributions of Earnings for Major Race-Sex Groups: A Dynamic Analysis", 73 *Am. Econ. Rev.,* 766 (1983); Charles Brown, Curtis Gilroy y Andrew Kohen, "The Effect of the Minimum Wage on Employment and Unemployment", 20 *J. Econ. Lit.,* 487 (1982); Daniel S. Hamermesh, "Minimum Wages and the Demand for Labor", 20 *Econ. Inquiry,* 365 (1982); Peter Linneman, "The Economic Impacts of Minimum Wages Laws: A New Look at an Old Question", 90 *J. Pol. Econ.,* 443 (1982); Jacob Mincer, "Unemployment Effects of Minimum Wages", 84 *J. Pol. Econ.,* S87 (1976); *The Economics of Legal Minimum Wages* (Simon Rottenberg [comp.], 1981). La mayoría de los economistas creen que el salario mínimo tiene ciertos efectos de desempleo por lo menos, pero algunos de ellos opinan de otro modo. Los opositores principales son David Card y Alan B. Krueger. Véase su libro *Myth and Measurement: The New Economics of the Minimum Wage* (1995); y una crítica de él en David Neumark y William Wascher, "Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania: Comment", 90 *Am. Econ. Rev.*, 1362 (2000); Richard V. Burkhauser, Kenneth A. Couch y David C. Wittenburg, "A Reassessment of the New Economics of the Minimum Wage Literature with Monthly Data from the Current Population Survey", 18 *J. Lab. Econ.*, 653 (2000); y Daniel Shaviro, "The Minimum Wage, the Earned Income Tax Credit, and Optimal Subsidy Policy", 64 *U. Chi. L. Rev.*, 405, 440-450 (1997). Adviértase que una ley de salario mínimo puede tener un efecto de desempleo secundario: al reducir el promedio de horas de trabajo. Véanse Thomas R. Michl, "Can Rescheduling Explain the New Jersey Minimum Wage Studies?", 26 *E. Econ. J.*, 265 (2000); Kenneth A. Couch y David C. Wittenburg, "The Response of Hours of Work to Increases in the Minimum Wage", 68 *S. Econ. J.*, 171 (2001). Sin embargo, los efectos de desempleo del salario mínimo se reducen por la violación generalizada de la ley del salario mínimo. Véase, por ejemplo, Yang-Ming Chang, "Noncompliance Behavior of Risk Averse Firms under the Minimum Wage Law", 20 *Pub. Finance Q.,* 390 (1992).



GRÁFICA XII.2

precio.) En virtud de que quienes ganan salarios bajos se encuentran a menudo en familias de altos ingresos, el salario mínimo no resulta ser una medida eficiente para combatir la pobreza, aun sin considerar sus efectos adversos sobre los trabajadores marginales.

Si el salario mínimo es aplicable sólo en algunas ocupaciones, podría reducir el desempleo en el sector exento y aumentarlo en los sectores cubiertos. La razón es que los trabajadores se verán atraídos al sector cubierto porque ahora los salarios son más altos allí, con lo que se reduce la oferta en el primero (y aumentan los salarios ahí) y aumenta en el segundo. Dependiendo de la diferencia salarial existente entre los sectores, los trabajadores podrían aceptar un periodo de desempleo prolongado en el sector cubierto como el precio de encontrar un empleo mejor pagado allí. Un camino para aguardar a que se ofrezca trabajo en el sector cubierto es el de la mayor permanencia en la escuela; por lo tanto, ¡el salario mínimo promueve la educación universitaria! ¿También aumenta la delincuencia porque el del delito es un sector exento? Eso es difícil de decir. El salario mínimo incrementa el índice de delincuencia para los trabajadores que están desempleados porque reduce las oportunidades legítimas de empleo y, desde luego, no hay salario mínimo en la delincuencia. Pero disminuye el índice de delincuencia de los trabajadores cuyos salarios se elevan por el salario mínimo, ya que los costos de oportunidad de la delincuencia aumentan para ellos.

Aunque los economistas conservadores consideran el salario mínimo como el ejemplo más claro (junto con el control de los alquileres y otros controles de precios) de la interferencia gubernamental gratuita con el libre mercado, la evidencia de que provoca desempleo significativo o tiene otros efectos negativos (adviértase en el párrafo anterior la naturaleza no concluyente del argumento de que el salario mínimo promueve la delincuencia) es en realidad bastante débil.[21] Es posible que los empleadores inviertan más en

[21] Véase, por ejemplo, Simon Deakin y Frank Wilkinson, *Minimum Wage Legislation, in Labor and Employment Law and Economics*, 150, 154-159 (Kenneth G. Dau-Shmidt, Seth D. Harris y Orly Lobel [comps.], 2009).

el capital humano de los trabajadores si tienen que pagarles más, aumentando la productividad, o que los trabajadores potenciales busquen más para conseguir mejores empleos, educación, etc., si están desempleados debido al salario mínimo, o que por debajo de cierto nivel salarial las personas prefieran la asistencia social, por lo que la asistencia tendría que reducirse por la abolición del salario mínimo para reducir el desempleo. Pero se podría producir una gran dificultad si, con beneficios mínimos de asistencia social y sin un salario mínimo, los trabajadores marginales se vieran obligados a aceptar empleos que sólo pagaran salarios de subsistencia. Se podría esperar que tal sistema produjera no sólo miseria, sino también delitos.

Muchas otras leyes, además de las leyes del salario mínimo, aumentan las tasas salariales de subconjuntos seleccionados de trabajadores. Las restricciones observadas en los códigos de la construcción para el empleo de materiales prefabricados aumentan la demanda de mano de obra según los lineamientos sugeridos en nuestra discusión de los acondicionadores de aire prefabricados. Las leyes del trabajo infantil se justifican incluso por razones económicas porque la mano de obra infantil tiende a destruir el capital humano (a los niños trabajadores de tiempo completo se les negaría una educación), pero al reducir la oferta de mano de obra esas leyes aumentan los salarios de los trabajadores adultos; las leyes "protectoras" de las mujeres trabajadoras tienen (o más bien tenían, porque ahora están extintas) un efecto similar y una justificación menor (véase después XXVI.1).

Los aranceles son frecuentemente buscados y obtenidos por los sindicatos como un método para contrarrestar la ventaja comparativa disfrutada por empresas extranjeras que emplean mano de obra más barata. Las leyes restrictivas de la inmigración también reducen la cantidad de mano de obra barata, pero su eficacia es limitada porque los empleadores se oponen a ellas y hasta ahora han tenido éxito para bloquear medidas efectivas de cumplimiento, mientras que los sindicatos son ambivalentes porque esperan organizar a los inmigrantes ilegales. La inmigración de mano de obra extranjera altamente calificada ha sido obstaculizada con mayor éxito porque se oponen a ella los trabajadores estadunidenses altamente calificados, quienes tienen más peso político que los no calificados.

La Ley de Seguridad y Salud en el Empleo (OSHA), que rige el establecimiento de estándares federales mínimos para la seguridad y la salud de los trabajadores, es un ejemplo particularmente ambicioso de la legislación protectora de los trabajadores, pero cabe tacharla de superflua. El empleador tiene un interés egoísta en proveer al trabajador el nivel óptimo (no necesariamente el más elevado posible) de salud y seguridad. Si pueden eliminarse por 99 centavos de dólar un accidente o una enfermedad esperados que cuestan 1 dólar, la eliminación reducirá la nómina salarial neta del empleador en un centavo, dado que sus empleados demandaron presumiblemente una compensación por el costo esperado de 1 dólar. La legislación que pres-

cribe las condiciones de salud y seguridad del empleo podría elevar el nivel de salud y seguridad por encima del nivel deseado por los empleados y los empleadores, y entonces ambos grupos resultarán perjudicados. Si la legislación le exige gastar 1.05 dólares en la eliminación de un riesgo de salud que imponga un costo esperado de sólo uno, el empleador reducirá la tasa salarial en 1 dólar por lo menos (ya no tendrá que compensar al empleado por el riesgo), y probablemente en 1.05 dólares. Cualquier aumento en el costo de un empleado reduce su producto marginal para la empresa y, por lo tanto, es probable que conduzca a una reducción correspondiente en su salario. Por eso son los empleados los que al final pagan no sólo la parte proporcional del impuesto a la seguridad social que se deduce directamente de sus salarios, sino también el porcentaje que el empleador paga nominalmente.

El análisis no cambia significativamente si suponemos que las tasas salariales nominales son inflexibles hacia abajo porque están especificadas en contratos de negociación colectiva, ya que estos contratos suelen expirar después de tres años. La próxima vez que el sindicato demande un aumento de 2 dólares en el salario por hora, por la inflación, el aumento de la productividad o lo que sea, sólo obtendrá 1 dólar, para compensar al empleador por haber eliminado el riesgo en la salud.

Pero es posible que todo esto sea demasiado cómodo y omita los costos de la información acerca de la seguridad en el lugar de trabajo. O bien es posible que la verdadera intención de la OSHA sea la disminución de la competencia de los trabajadores no sindicalizados (aunque la OSHA también podría afectar a los sindicatos porque, en su ausencia, el papel de los sindicatos en la negociación de condiciones de trabajo seguras sería más importante, y esto haría que los sindicatos fueran más atractivos para los trabajadores con ocupaciones peligrosas). Si el nivel de la salud y la seguridad es generalmente menor en las empresas que pagan salarios no sindicales —un supuesto plausible porque entre menor sea el salario del trabajador menores serán para él los costos de verse incapacitado—, los sindicatos podrían presionar por una legislación que eleve el nivel de la salud y la seguridad en esas empresas hasta el nivel prevaleciente en las empresas sindicalizadas. Los empleadores no sindicalizados podrían tratar de compensar el aumento de sus costos reduciendo los salarios que pagan a sus trabajadores. Pero, si tales trabajadores estuvieran ganando ya apenas el salario mínimo, los empleadores no podrían hacer ninguna reducción salarial compensatoria, y los costos laborales aumentarían en relación con los de las empresas sindicalizadas. Esta hipótesis es apoyada por la preferencia de la Administración de Seguridad y Salud en el Empleo, aparentemente motivada por el deseo de minimizar el efecto de sus regulaciones sobre el empleo, de instrumentar cambios en los métodos de trabajo destinados a reducir los accidentes en el lugar de trabajo en vez de depender de artículos de protección personal tales como tapones auriculares y anteojos de seguridad. Es menos probable que los cambios en los métodos

aumenten el costo marginal de la mano de obra; una vez que se realiza tal cambio, puede aumentar la seguridad del lugar de trabajo independientemente del número de trabajadores, mientras que el equipo de seguridad individual aumenta el costo de cada empleado para el empleador. Por lo tanto, es menos probable que los cambios en los métodos de trabajo tengan un efecto significativo en el desempleo. A pesar de que los dispositivos de protección personal suelen ser más eficientes que los cambios en los métodos de trabajo (como lo indican las pruebas existentes en el sentido de que, aunque la OSHA aumenta indudablemente los costos de los empleadores, no ha disminuido significativamente el número de las lesiones en el lugar de trabajo[22]), los cambios en los métodos aumentan el costo del capital y los dispositivos de protección personal incrementan el costo de la mano de obra, y de esta manera reducen los salarios o el empleo o ambos.[23]

Pero los trabajadores de plantas no sindicalizadas tienen en realidad menos información acerca de la seguridad, porque un sindicato es un recolector de información para sus miembros. La OSHA corrige las ineficiencias resultantes de programas gubernamentales en apoyo de los trabajadores lesionados. Mediante el programa de incapacidad de la seguridad social (véase después XVII.6), además de diversos programas especiales (por ejemplo, el programa de compensación de los trabajadores incapacitados por la enfermedad de los pulmones negros), el gobierno federal paga una parte del costo en que incurren los trabajadores lesionados a resultas de condiciones de trabajo inseguras o insalubres. Trasladando los costos del accidente y la enfermedad de los trabajadores y sus empleadores al contribuyente federal, esos programas reducen los incentivos de las empresas para mantener el nivel óptimo de la seguridad y la salud ocupacionales.

Sin embargo, utilizar una intervención gubernamental en el mercado (subsidiando las lesiones y enfermedades en el lugar de trabajo) a fin de justificar otra (la regulación de la seguridad y la salud ocupacionales) invita a una expansión indefinida del gobierno. Una serie de pasos incrementales, cada uno de los cuales tiene sentido económico a la luz de los pasos previos, podría no tener ningún sentido económico si se observa como un todo.

[22] Ann P. Bartel y Lacy Glenn Thomas, "Direct and Indirect Effects of Regulation: A New Look at OSHA's Impact", 28 *J. Law & Econ.*, 1 (1985); David P. McCaffrey, "An Assessment of OSHA's Recent Effects on Injury Rates", 18 *J. Human Resources*, 131 (1983); W. Kip Viscusi, *Risk by Choice: Regulating Health and Safety in the Workplace*, cap. 2 (1983); Viscusi, "The Impact of Occupational Safety and Health Regulation", 10 *Bell J. Econ.*, 117 (1979). Pero véase John T. Scholz y Wayne B. Gray, "OSHA Enforcement and Workplace Injuries: A Behavioral Approach to Risk Assessment", 3 *J. Risk & Uncertainty*, 283 (1990).

[23] James C. Miller III, "Is Organized Labor Rational in Supporting OSHA?", 50 S. *Econ. J.*, 881 (1984).

XII.9. Beneficios ordenados

Algunos estados de la Unión Americana exigen ahora que los empleadores provean beneficios de maternidad a sus empleados. Un estudio ha descubierto que este requerimiento hace que los salarios bajen por el costo total de los beneficios para los empleadores.[24] Puede verse fácilmente por qué bajarían los salarios; los costos laborales de los empleadores han aumentado. Si los salarios no bajaran, el otorgamiento de este beneficio accesorio sería el equivalente de un aumento; y si los empleadores desearan otorgar a sus empleados un aumento bajo la forma de beneficios de maternidad, ¿por qué esperarían a la promulgación de la ley? Pero los salarios no debieran bajar en el monto *total* del aumento del costo para los empleadores, porque eso implicaría que el valor del nuevo beneficio no era mayor a su costo, ¿y entonces por qué los empleadores no habrían ofrecido el nuevo beneficio voluntariamente? Además, dado que no todos los empleados se proponen tener hijos, los beneficios de la maternidad no valen la misma cantidad para todos los empleados; en efecto, no valen nada para muchos de ellos.

¿Cómo explicar entonces la anomalía de que los salarios bajen por el monto total del costo del beneficio recién ordenado? Una posibilidad es el efecto de la dotación, una de las peculiaridades cognitivas iluminada por la psicología: es posible que los trabajadores no hayan deseado el beneficio pero que, una vez obtenido, lo evaluaran altamente porque era suyo, de modo que estaban dispuestos a pagar por el beneficio en su totalidad bajo la forma de una reducción de sus salarios. (Es cierto, los salarios también eran "suyos"; entonces, ¿por qué estaban dispuestos a renunciar a una parte de ellos? Tal vez porque no se les dio la opción, o porque la pérdida de salarios podría tomar la forma, que puede ser invisible para la mayoría de los trabajadores, de un aumento más lento de los salarios en lugar de una reducción real.) En segundo lugar, podría haber un efecto de selección. Los trabajadores que no deseen los beneficios de la maternidad tenderán a trasladarse a estados donde tales beneficios no formen parte del paquete de compensación, y serán remplazados por trabajadores que valúen grandemente tales beneficios.

XII.10. Discriminación en el empleo por la raza, el género, la edad y la discapacidad

La discriminación racial en el empleo forma parte de una cuestión más amplia —la de las causas y los remedios de la discriminación racial— que se

[24] Jonathan Gruber, "The Incidence of Mandated Maternity Benefits", 84 *Am. Econ. Rev.,* 622 (1994).

discute en un capítulo posterior (el capítulo XXVIII). Aquí examinaremos un tema especializado de la discriminación racial, así como la discriminación en el empleo contra los trabajadores sobre la base del género, la edad o la incapacidad.

1. Muchos sindicatos se negaron durante largo tiempo a admitir trabajadores negros. Como hemos visto, los sindicatos tratan de aumentar la tasa salarial por encima del nivel competitivo, y, en la medida en que lo logren, se generará una demanda excedente de empleos sindicalizados. Esta demanda excedente podría eliminarse en diversas formas. Una de ellas sería la subasta de la membresía de empleados sindicalizados. Los postores triunfantes serían quienes estuvieran dispuestos a pagar una cuota de entrada igual al valor presente de la diferencia existente entre la escala salarial sindical y los salarios de su siguiente mejor empleo. Éste sería el método de racionamiento que se emplearía si los sindicatos fueran simplemente empresas que disfrutaran un poder monopólico sobre los trabajadores: empresas que compraran trabajadores al salario competitivo y los revendieran a los empleadores a un salario monopólico. Pero los sindicatos no son empresas; son representantes, por imperfectos que sean, de los trabajadores. No adoptarán un método de racionamiento que negara a los trabajadores cualquier ganancia neta de su calidad de miembro de un sindicato.

No obstante, el problema de los métodos de racionamiento no monetarios es que inducen un gasto de recursos reales. Por ejemplo, si se utilizara el método de las colas —los solicitantes potenciales de la calidad de miembro de un sindicato se pasan años como aprendices mal pagados, esperando que surja un empleo—, los ingresos de ocupaciones alternativas sacrificados se convertirían en el costo de la membresía y podrían disipar las ganancias de ésta. Lo que vuelve atractivos para los sindicatos a los criterios que involucran la raza o algún otro estatus inmutable (como el hecho de ser hijo de un miembro del sindicato) es que no implican grandes gastos para la calificación; normalmente son infinitos los costos del cambio de raza o de padres.[25]

2. La cuestión económica central en lo tocante a la discriminación contra las mujeres en el empleo consiste en explicar el salario medio persistentemente mayor de los hombres; en promedio, los salarios por hora de las mujeres son iguales a tres cuartos de los salarios de los hombres. Las diferencias existentes entre las inversiones hechas en el capital humano relacionado con el mercado parecen ser la clave para entender este fenómeno. La mujer que dedica una parte sustancial de su vida de trabajo a la producción hogareña, incluida la

[25] Sin duda, la discriminación racial por parte de los sindicatos también reflejaba otros factores. Uno de ellos era que los sindicatos querían lograr la "respetabilidad" (la cual tendría valor político) al disociarse de un grupo que, en el pasado reciente de entonces, había realizado trabajo esclavo y tenía un estatus social muy bajo. Además, un cártel es más unido mientras más homogéneos socialmente sean sus miembros, porque eso tiende a facilitar la comunicación y crear confianza.

atención de los hijos, obtendrá un rendimiento total sustancialmente menor sobre su capital humano de mercado que un hombre que dedica mucho menos tiempo a tal producción. Por lo tanto, ella invertirá menos en ese capital humano de mercado, y, dado que los ingresos son en parte un rendimiento de nuestro capital humano, los ingresos de las mujeres serán menores que los de los hombres.[26] Por lo tanto, las mujeres se verán atraídas hacia las ocupaciones que no requieren mucho capital humano; ésta es una de las razones por las que las mujeres tendían a ser enfermeras antes que médicos, y secretarias antes que ejecutivos. Pero, en la medida en que las mujeres dedican más tiempo que los hombres a la producción hogareña porque las oportunidades para las mujeres en el mercado de empleos están deprimidas por la discriminación de género, la diferencia existente en la cantidad de capital humano de mercado que acumulan los hombres y las mujeres y la diferencia resultante de los ingresos son un artificio de la discriminación, pero discriminación en toda la sociedad, no una discriminación de la que pueda culparse a un empleador particular. Sin embargo, dado que la cantidad de tiempo que las mujeres están dedicando a la producción hogareña está disminuyendo por las razones discutidas en el capítulo v, las diferencias ocupacionales y salariales existentes entre hombres y mujeres se habrían reducido aun sin la intervención gubernamental. La existencia de la discriminación por género de parte de los empleadores es una condición necesaria, antes que suficiente, para que una ley que prohíba tal discriminación pueda conferir un beneficio económico neto a la sociedad en conjunto, o aun al conjunto de las mujeres. Primero, no toda discriminación por razones de género es ineficiente. Considérese la negativa (que ahora es ilegal) de los empleadores a pagar beneficios de incapacidad por embarazo. Desde el punto de vista de la determinación eficiente de los precios de la mano de obra, obligar a que se pague a una trabajadora que está ausente por causa de embarazo no tiene más sentido que obligar a que se pague a un trabajador que está ausente porque está pintando el cuarto del bebé. En ningún caso está el trabajador confiriendo un beneficio al empleador. (¿En qué difieren las vacaciones pagadas o las ausencias por enfermedad pagadas? Continúe leyendo.)

Segundo, las leyes contra la discriminación pueden revertirse contra la clase protegida cuando los empleadores toman medidas racionales para minimizar el impacto de las leyes. Por ejemplo, la Ley de Paga Igual ordena a los empleadores que paguen el mismo salario a hombres y mujeres que estén realizando el mismo trabajo. Si por cualquier razón las mujeres trabajadoras tienen un producto marginal menor, los empleadores tendrán un incentivo para sustituir insumos de mano de obra por insumos de capital en las clasifi-

[26] Jacob Mincer y Haim Ofek, "Interrupted Work Careers: Depreciation and Restoration of Human Capital", 17 *J. Human Resources,* 3 (1982); Mincer y Solomon W. Polachek, "Family Investments in Human Capital: Earnings of Women", 82 *J. Pol. Econ.,* S76 (1974); "Trends in Women's Work, Education, and Family Building", 3 *J. Labor Econ.,* S1 (1985).

caciones de empleos que ocupen a muchas mujeres. Además, el hecho de tener que pagar a las empleadas más que su valor de mercado podría hacer que los empleadores se resistieran a contratar mujeres, y aunque esta renuencia podría exponer a un empleador a una demanda judicial por discriminación en la contratación, tales demandas son raras porque los daños esperados suelen ser pequeños. La razón es que es improbable que una solicitante de empleo acabe obteniendo un trabajo que pague mucho más de lo que podría obtener en otra parte, pero, si su solicitud es rechazada injustamente, los daños se limitarán a la diferencia existente entre el salario que la demandante habría obtenido en el empleo del empleador discriminador y el que obtiene en su empleo actual o, si no está empleada, entonces en su mejor empleo alternativo.

Aquí existe un doble revés. Mientras que los costos de responsabilidad esperados de una negativa de contratación discriminatoria son bajos, los costos de responsabilidad esperados de un despido discriminatorio (llamémoslos "costos de despido") son altos, ya que es probable que el salario de un trabajador en su trabajo actual, debido a que tiene capital humano específico de la empresa, supere, a menudo por un margen considerable, el salario en el mejor trabajo alternativo que pudiera obtener. Estos costos se evitan al no contratar al miembro de una clase protegida por la ley de discriminación laboral.

Incluso si la Ley de Paga Igual se aplica plenamente, las mujeres podrían no beneficiarse. Tener que pagarles un salario mayor aumentará los costos del empleador y, por lo tanto, puede hacer que éste disminuya los salarios de todos sus empleados, masculinos y femeninos, o aumente esos salarios más lentamente de lo que lo haría, o despida empleados. Entre mayor sea la fracción de la fuerza de trabajo femenina, menores serán los costos añadidos del empleador que se exportarán de estas maneras a los trabajadores masculinos y, por lo tanto, será menor el beneficio que obtengan las mujeres de la ley.[27]

Además, algunos de los costos que las leyes de antidiscriminación imponen a los empleadores tenderán a trasladarse a los consumidores bajo la forma de precios mayores, y los consumidores femeninos se verán perjudicados junto con los masculinos. La heterogeneidad de los intereses de las mujeres, combinada con la interdependencia financiera y altruista existente entre hombres y mujeres, vuelve más incierto aún que las mujeres sean beneficiarias netas de las leyes de antidiscriminación. El aumento en los costos laborales derivados del pago de beneficios de la incapacidad por embarazo son sufragados, en parte, por todos los trabajadores, pero mientras que las empleadas embarazadas obtienen una compensación de los mismos beneficios, las mujeres que no están empleadas o que no son fértiles serán claras perdedoras.

Los mayores perdedores son las mujeres no empleadas o que no planean tener hijos pero que estén casadas con hombres empleados. Los salarios o

[27] Véase Christine Jolls, "Antidiscrimination and Accommodation", 115 *Harv. L. Rev.,* 642 (2001).

las oportunidades de empleo de sus esposos disminuirán como resultado de los costos de mano de obra más altos de los empleadores causados por los beneficios obligatorios de los empleados, y las esposas sufrirán en la medida en que el consumo hogareño sea conjunto o en que ellas sean altruistas hacia sus esposos. Además, todos los consumidores tendrán que pagar precios mayores si los insumos de mano de obra se emplean menos eficientemente a resultas de que se declare ilegal la discriminación eficiente.

El "valor comparable" se refiere a la política, cuyos proponentes creen que debiera estar inscrita en la ley, de igualar los niveles salariales en los empleos ocupados primordialmente por mujeres y en los empleos dominados por los hombres determinando el valor "efectivo" de empleos diferentes y requiriendo que los empleos que tengan el mismo valor efectivo involucren el mismo salario. Si se prohíbe que los empleadores excluyan a las mujeres de las clasificaciones de empleos deseadas, el valor comparable carece de sentido económico. Porque, según ese supuesto, si se le paga más a un camionero que a una secretaria, aunque la secretaria trabaje una jornada de igual duración y tenga una educación igualmente buena, el mercado debe de estar compensando una habilidad cuya oferta es menor, o contrarrestando alguna inconveniencia, antes que haciendo distinciones arbitrarias basadas en estereotipos que se están esfumando rápidamente.

Según el valor comparable, si se elevan artificialmente los salarios en los empleos que están dominados por las mujeres, disminuirá el número de los empleos disponibles porque los empleadores tratarán de emplear otros insumos que ahora son más baratos, mientras que los consumidores comprarán otros productos en lugar de los que fabrican las empresas cuyas nóminas, y por lo tanto sus precios, han aumentado a causa del valor comparable. Al mismo tiempo, los hombres empezarán a competir más por tales empleos, atraídos por los salarios mayores. Por lo tanto, el empleo femenino en una clasificación que había sido propia de las mujeres podría disminuir. Algunas mujeres desplazadas encontrarán un nuevo empleo en las ocupaciones predominantemente masculinas, como la de los camioneros… quizá remplazando a hombres que se han convertido en secretarios. Sin embargo, algunas mujeres que solían trabajar en "áreas masculinas" ahora encontrarán empleos "femeninos" más atractivos debido al cambio en las tasas salariales relativas. De modo que, mientras que algunas mujeres huirán de las clasificaciones de empleo "femenino", otras se desplazarán hacia ellas en masa. Por último, donde se aplique el valor comparable disminuirán los incentivos de las mujeres para invertir en capital humano utilizable en los empleos tradicionales de los hombres, a medida que disminuyan los salarios relativos de tales empleos, de modo que en última instancia la segregación ocupacional por el género no se vería muy afectada.

3. La Ley de Discriminación en el Empleo por la Edad prohíbe que los empleadores públicos y privados fijen una edad de jubilación obligatoria,

con la excepción de unas cuantas clasificaciones de empleos tales como la del piloto de avión y el ejecutivo empresarial de alto nivel, o que se discrimine con base en la edad. El argumento de la ley es que la utilización de la edad como un criterio para el empleo es arbitraria porque la gente envejece a ritmos diferentes. Esto no provee una buena razón para el derecho. La utilización de una característica única, fácilmente determinable tal como la edad para que sirva de base a la decisión sobre el empleo economiza los costos de la información. El hecho de que a veces sea remplazado un trabajador anciano más competente por un trabajador más joven pero menos competente no vuelve ineficiente la utilización de la edad como una aproximación a la competencia, por cruda que sea. El objetivo del empleador es minimizar la suma de los costos de las decisiones de retención subóptimas resultantes de la ausencia de evaluaciones individualizadas de la capacidad de los trabajadores y los costos de información de la realización de tales evaluaciones.[28] Si se minimiza la suma estableciendo una edad de jubilación obligatoria, el empleador fijará tal edad; de otro modo no lo hará. No hay ninguna externalidad que amerite la intervención gubernamental.

Y si son elevados los costos de las evaluaciones individualizadas de las capacidades de los trabajadores (y si no lo fuesen, ¿para qué utilizaría el empleador la edad como un discriminante?; ¿qué tan probable es que las personas de edad mediana que dirigen las corporaciones tengan una hostilidad irracional hacia las personas de su propia edad?), entonces al prohibir a los empleadores que utilicen la edad como una aproximación utilizarán otras aproximaciones, tales como los resultados de exámenes, discriminando así contra los trabajadores cuyo desempeño es pronosticado deficientemente por tales aproximaciones. Otros empleadores se encogerán de hombros, incapaces para distinguir entre los trabajadores buenos y los malos, y tratarán a ambos grupos indiscriminadamente, con el resultado de que los trabajadores malos se beneficiarán a expensas de los buenos.

Un argumento mejor a favor de la ley dice que si los trabajadores excepcionales de mayor edad —los que son jóvenes de mente, cuerpo y espíritu— no pueden hacer valer su excepcionalidad en el mercado de empleos porque muy pocos empleadores verán más allá de la edad cronológica al tomar sus decisiones de empleo, estos trabajadores tendrán un incentivo subóptimo para invertir en su capital humano porque se truncará artificialmente el periodo de la recuperación. Un trabajador individual podría, aceptando un salario menor, inducir al empleador escéptico a hacer una excepción en su favor. Pero la perspectiva del salario menor reduciría el rendimiento esperado de la inversión del empleado excepcional en su capital humano y, por ende, el monto de lo que estaría dispuesto a invertir.

[28] Este problema decisorio se asemeja al de decidir cuánto procedimiento conviene proveer a los litigantes (más adelante, XXII.1), o si convendrá utilizar una regla o un estándar para gobernar la responsabilidad (después, XXI.3).

No obstante, si los costos de las evaluaciones individualizadas del desempeño de los empleados son tan grandes que prohibir la utilización de la aproximación de la edad conduce al empleo de otras aproximaciones, la ley que prohíbe la discriminación por la edad no resolverá el problema de la inversión deficiente. Quienquiera que se vea injustamente perjudicado por la nueva aproximación carecerá de incentivo para hacer la inversión óptima en su capital humano. Y, si las nuevas aproximaciones son menos eficientes que la edad —como ocurrirá probablemente, porque de otro modo se habrían adoptado sin la intervención gubernamental—, los salarios bajarán porque los costos laborales de los empleadores serán mayores.

El mejor argumento en favor de la ley que prohíbe la discriminación por la edad es que constituye un método para el cumplimiento obligatorio de un contrato implícito por el que un empleador consiente en pagar un salario mayor al producto marginal del empleado a fin de compensarlo por invertir en capital humano específico de la empresa. Supongamos que el trabajador pagara por el grueso de ese capital aceptando un salario menor al del mercado en sus primeros años, esperando recuperarse más tarde con un salario mayor al del mercado. El empleador podrá frustrar esta recuperación despidiéndolo antes de tiempo. Esto no tendría sentido si, a resultas de la inversión de la empresa en su capital humano, el trabajador fuese tanto más productivo al salario mayor que resultase más valioso para la empresa a ese salario de lo que sería un trabajador joven a un salario menor. Pero, después de cierto punto, la productividad de un trabajador suele disminuir con la edad, aunque podría encontrarse todavía en su periodo de recuperación, ganando un salario que podría exceder a su producto marginal actual para normalizar sus salarios a lo largo de su vida de trabajo. El empleador se verá tentado a despedirlo a menos, desde luego, que lo detengan consideraciones de la reputación o un contrato de trabajo.

O bien es posible que el empleador haya establecido una disposición salarial diferida a fin de asegurar el buen desempeño del empleado. Supongamos que es la clase de empleo en el que un empleado puede calumniar o aun robar a su empleador, y que los costos de la detección son elevados. A fin de aumentar el costo esperado, para el empleado, de ser detectado y despedido —en otras palabras, a fin de aumentar el costo esperado del castigo (*PL*) sin aumentar *P,* lo que significa aumentar *L*—, el empleador podría desplazar la remuneración hacia el final de la carrera del empleado, por ejemplo, como veremos en la sección siguiente, bajo la forma de una pensión generosa.[29]

[29] El argumento más general es que "pagar de más" a un agente (es decir, pagarle una renta económica, o sea, una remuneración por encima de sus costos de oportunidad) es un método para reducir la flojera u otro mal comportamiento del agente aumentando el costo del despido para él. Por ejemplo, un fabricante podría tratar de fijar un precio de reventa que les rindiera cierta renta económica a sus distribuidores, de modo que resulte más eficaz la amenaza de despedirlos por falta de celo o de honestidad. *Cf*. antes X.4.

(Ésta es una versión de la teoría económica del castigo [véase el capítulo VII], en la que el castigo se hace más severo para compensar una probabilidad mínima de ser impuesto.) Con todo, una vez que el empleado se encuentra más cerca del final de su vida de trabajo, el empleador tendrá un incentivo para despedir al empleado si de este modo puede liberarse de la obligación de la compensación diferida. Desde luego, si los empleadores observan la clase de comportamiento oportunista que acabamos de describir, los empleados podrían demandar contratos que prohíban la terminación sin causa justa, tal como lo determine un árbitro u otro juzgador neutral, o exigir salarios mayores en las etapas tempranas de su carrera para compensarlos por el riesgo de mandarlos a su casa antes de tiempo. Aunque la reputación por sí sola no restringirá a los empleadores en su último periodo, el empleado tendrá poco que perder por el despido: si el empleador está a punto de abandonar la industria, sus empleados serán despedidos pronto, cualquiera que sea su edad.

Una razón para dudar de la eficacia de la discriminación por la edad es que tanto la prohibición de la jubilación obligatoria a una edad fija como la prohibición más amplia de la discriminación en el empleo basada en la edad pueden eludirse mediante ofrecimientos de jubilación temprana. A pesar de las apariencias, la oferta de jubilación temprana como un procedimiento para reducir la edad media de la fuerza de trabajo de un empleador es un instrumento de zanahoria y garrote, no sólo una zanahoria… y la parte de la zanahoria resulta muy barata para el empleador, ya que en efecto la pagan los propios empleados. El derecho a aceptar la jubilación temprana en condiciones suficientemente ventajosas para que resulte atractivo el ejercicio del derecho es una forma de compensación del empleado que no resulta diferente de cualquiera otra prestación accesoria. Su valor es incierto, pero lo mismo ocurre con el valor del seguro médico o de vida, que depende para un trabajador individual de su experiencia particular de salud y de longevidad, cuyo valor no puede determinarse por adelantado con certeza. Entre más generoso sea el ofrecimiento de la jubilación temprana, menos tendrá que pagar el empleador por concepto de salarios y otras prestaciones.[30]

El costo neto para el empleador del ofrecimiento de una jubilación temprana está restringido también por el riesgo que corre un empleado de rechazar incluso un ofrecimiento mediocre por tacañería; aquí es donde entra en juego

[30] La creación de un programa generoso de jubilación temprana financiado (en equilibrio) por una reducción del nivel salarial volvería al empleador más atractivo para los trabajadores que otorgan gran valor al ocio en relación con el ingreso pecuniario, mientras que el empleador podría preferir a trabajadores con una ética de trabajo más sólida. Pero entre mayor sea el número de los empleadores que adopten tales programas, menor será el efecto sobre la composición de cualquier fuerza de trabajo dada del empleador.

Adviértase que, aun si no previó la promulgación de la ley contra la discriminación por la edad, los trabajadores cercanos a la edad de jubilación obtuvieron de esa ley un beneficio neto esperado. Eso no ocurrió con los trabajadores jóvenes, porque sus salarios se ajustarían para contrarrestar los beneficios que esperaban de la ley cuando fuesen más viejos.

el garrote. A menos que el empleado pueda probar que el paquete de jubilación y otras prestaciones ofrecido por el empleador no es un plan de beneficios de buena fe, sino que está diseñado para evadir la ley que prohíbe la discriminación por la edad (algo que no resulta fácil probar), el empleador puede castigar al empleado por rehusar la oferta de una jubilación temprana ofreciendo en cambio prestaciones menores para los empleados que se jubilen más tarde. Además, el ofrecimiento de una jubilación temprana no compromete al empleador a retener hasta la edad normal de la jubilación a un empleado que rechace la oferta. Por el contrario, tal oferta refleja de ordinario el deseo del empleador de reducir el número o la edad promedio de sus empleados, y si la oferta no es aceptada por muchos empleados, y aunque muchos la acepten, el empleador podrá tomar otras medidas para obtener el tamaño y la composición por la edad deseados de su fuerza de trabajo. El empleado lo sabe, de modo que está consciente de que si rechaza hoy una oferta de jubilación temprana podría ser despedido mañana en circunstancias que podrían hacerle imposible probar la discriminación por la edad y obtener así una compensación (porque el empleador podrá demostrar la necesidad que tiene la empresa de reducir su fuerza de trabajo) o encontrar otro empleo con un salario equivalente, sobre todo si su salario actual incluye una compensación por el capital humano específico de la empresa.

Cuanto más viejo sea un trabajador, es más probable que los "costos de despido" del empleador sean mayores. Por un lado, esto hace que sea menos probable que el trabajador más viejo sea despedido, pero por otro lado hace más probable que no sea contratado en primer lugar. Este último efecto, junto con la estrategia de jubilación temprana para reducir la edad media de la fuerza de trabajo, podría explicar por qué se ha determinado que la Ley de Discriminación en el Empleo por la Edad tiene un efecto negativo neto sobre el empleo de los trabajadores de más edad.[31]

4. La Ley de los Estadunidenses con Discapacidades (ADA) contiene una prohibición contra la discriminación que afecte a personas con discapacidad que, a pesar de ésta, puedan desempeñarse con las especificaciones del empleador si éste "acomoda" la discapacidad de la persona, por ejemplo, proveyendo acceso para sillas de ruedas. El objetivo es aumentar las oportunidades de empleo para personas discapacitadas; sin embargo, el efecto real de la ley ha sido la reducción del número de personas discapacitadas que están empleadas.[32] Por otra parte, como otras leyes que combaten la discriminación, ADA

[31] Véase, por ejemplo, Joanna Lahey, "State Age Protection Laws and the Age Discrimination in Employment Act", 51 *J. Law & Econ.,* 3 (2008). Se han encontrado resultados similares con respecto a las leyes estatales que limitan el empleo a voluntad. Véase Max Schanzenbach, "Exceptions to Employment at Will: Raising Firing Costs or Enforcing Life-Cycle Contracts?", 5 *Am. Law & Econ. Rev.*, 470 (2003).

[32] Véanse, por ejemplo, *The Decline in Employment of People with Disabilities: A Policy Puzzle* (David C. Stapleton y Richard V. Burkhauser [comps.], 2003); Daron Acemoglu y Joshua D. An-

actúa como un subsidio a la contratación de trabajadores discapacitados, porque la negativa de contratar a una persona discapacitada genera el riesgo de una demanda judicial, lo que no ocurre con la negativa a la contratación de una persona no discapacitada. Ante una elección entre solicitantes igualmente adecuados, al empleador le resultará más barato contratar al solicitante discapacitado. Por otro lado, y aparentemente contrarrestando con creces el subsidio a la contratación, hay costos del acomodo de trabajadores incapacitados y los costos de litigación esperados si el trabajador discapacitado resulta poco satisfactorio y debe ser despedido, e interpone entonces una demanda en la que alega que fue despedido en violación de la Ley. No se permite que el empleador contrarreste estos costos ofreciéndole al trabajador discapacitado un salario menor al del trabajador no discapacitado que realiza el mismo trabajo.

La Ley impone tanto costos de despido (como otras leyes que prohíben la discriminación laboral que hemos examinado) como costos de acomodo, aunque un estudio revela que sólo estos últimos tienen un efecto de desempleo.[33] La obligación de acomodar la discapacidad se asemeja al enfoque del "efecto disparejo"[34] para probar la discriminación. Supongamos que un empleador requiere que todos sus empleados midan 1.70 metros por lo menos. Tal requerimiento tendría un efecto desproporcionado sobre el empleo de las mujeres. Aun cuando no pudiera probarse que el empleador hubiese adoptado el requerimiento como una aproximación al del género, el requerimiento sería ilegal a menos que el empleador pudiera probar que era una necesidad auténtica de su empresa. Esta carga de la prueba volvería más costoso para el empleador dar ocupación a mujeres (suponiendo que el requerimiento de la estatura tuviera alguna relación con la productividad, aunque no fuese suficiente para satisfacer la carga de la prueba del empleador), del mismo modo que ocurriría con la necesidad de construir instalaciones especiales para las mujeres.

Sin embargo, una diferencia entre el acomodo y el efecto disparejo es que a un empleador le resultaría más fácil sustituir con una acción afirmativa al último que al primero. Si, en el ejemplo de la estatura, mediante un reclutamiento activo, se las arreglara el empleador para contratar a suficientes mujeres de 1.70 para descartar toda inferencia de discriminación, podría conservar el requerimiento de la estatura, y los beneficios de la productividad podrían superar al costo del reclutamiento adicional de mujeres.[35]

grist, "Consequences of Employment Protection? The Case of the Americans with Disabilities Act", 109 *J. Pol. Econ.*, 915 (2001).

[33] Christine Jolls y J. J. Prescott, "Disaggregating Employment Protection: The Case of Disability Discrimination" (Harvard Public Law Wkg. Paper No. 106, 2005).

[34] Véase Jolls, nota 27 anterior.

[35] Shelly J. Lundberg, "The Enforcement of Equal Opportunity Laws under Imperfect Information: Affirmative Action and Alternatives", 106 *Q. J. Econ.*, 309 (1991). Adviértase que, si el requerimiento fuese el de tener un diploma del bachillerato o de la universidad o el de obtener cierta calificación en un examen, el efecto de la acción afirmativa (por ejemplo, exentar el reque-

XII.11. La Ley Federal de Pensiones

Existen dos tipos básicos de planes de pensión. Un plan de beneficios definidos promete pagar al empleado una pensión anual fija basada en el salario percibido durante su último año o (más comúnmente) sus últimos años de trabajo y en sus años de servicio. Entre mayor sea su salario a la terminación, y más prolongado el tiempo que haya trabajado para este empleador, mayor será la pensión a la que tiene derecho; y esto es independiente del monto de las contribuciones del empleador al plan o del desempeño de la inversión del plan.[36] En un plan de contribución definida, el empleador y el empleado suelen contribuir a una cuenta separada para cada empleado cubierto.[37] Cuando el empleado se jubila, la cantidad de dinero que haya en su cuenta, que será la suma de sus contribuciones y las de su empleador más cualquier interés u otro rendimiento ganado sobre la inversión de tales contribuciones, se usará para comprar una anualidad que le otorgue a él y a su esposa un ingreso asegurado por el resto de sus vidas. Así pues, el riesgo de la inversión en un plan de beneficios definidos corre a cargo del empleador o del plan (o del asegurador del plan), mientras que en un plan de contribuciones definidas corre a cargo del empleado, aunque éste puede eliminar la mayor parte del riesgo invirtiendo las contribuciones en una cartera de valores convenientemente diversificada (véase el capítulo XVI).

En virtud de que los beneficios, en un plan de beneficios definidos, se ligan al sueldo final del empleado, tal plan provee una seguridad mayor del ingreso salarial estable, aunque de ordinario no un remplazo completo, de modo que el nivel de vida acostumbrado del empleado será preservado aproximadamente en el retiro. Bajo un plan de contribuciones definidas, los beneficios dependen de la cantidad aportada y del desempeño de la inversión de las contribuciones. Si el empleado es el mejor asumidor del riesgo, la forma de los beneficios definidos será generalmente más atractiva para los empleados que la forma de las contribuciones definidas, aunque las diferencias existen-

rimiento para el grupo favorecido) sería una reducción del incentivo de los miembros del grupo para obtener educación por comparación con el incentivo de los miembros de la mayoría desfavorecida.

[36] Por ejemplo, la pensión anual podría ser 2% del salario final del empleado multiplicado por el número de sus años de servicio, de modo que si se retira después de 30 años tendría derecho a una pensión igual a 60% de su salario final.

[37] En el formulario de "saldo en efectivo" del plan de beneficios definidos, el empleado tiene una cuenta de jubilación "nocional" (nocional porque la cuenta individual no está financiada) a la cual cada año el empleador agrega un porcentaje específico del salario del empleado más los intereses a una tasa especificada sobre la cantidad en la cuenta. Si el empleado permanece empleado por el patrocinador del plan hasta la edad de retiro, el saldo de su cuenta de jubilación nocional se utilizará para comprar una anualidad para él o, si lo prefiere, podrá recibir el saldo en efectivo como una suma global.

tes en las reglas del otorgamiento y el riesgo omnipresente de la inflación complican la comparación de pérdidas sufridas y ganancias cosechadas. El plan de beneficios definidos facilita también (como veremos enseguida) las inversiones óptimas en capital humano, y confiere ventajas fiscales al empleador cuando el plan está financiado en exceso —es decir, cuando sus activos superan a los pasivos de la pensión—. El ingreso proveniente de tales activos irá en parte a las manos del empleador, porque no se necesitará todo el ingreso para el pago de los pasivos de la pensión. Ese ingreso se acumulará también libre de impuestos porque el plan es una entidad exenta de impuestos.

Antes de la promulgación de ERISA,[38] los pasivos de la pensión podían ser, y a menudo eran, pasivos del plan de pensiones solamente, antes que también del empleador que había creado el plan. Si el plan estaba financiado deficientemente porque las contribuciones habían sido inadecuadas, el desempeño de la inversión decepcionante o las proyecciones actuariales incorrectas, la pérdida recaía en los empleados; en esta forma, el riesgo de la inversión se trasladaba en parte a estos últimos. Si el plan estaba financiado en exceso, los accionistas del empleador, siendo los propietarios residuales de los fondos del plan después de haber cubierto todas las reclamaciones de los empleados, recogían la ganancia. Dependiendo de las reglas del otorgamiento y las de acreditación de los años de servicio adoptadas por el plan particular, un trabajador que se marchara antes de la edad de jubilación podría encontrarse con una pensión escasa, o incluso quizá con nada en absoluto. Por lo tanto, el trabajador tenía un fuerte incentivo para permanecer en la misma compañía hasta que alcanzara la edad de retiro. Este incentivo reducía la movilidad de los trabajadores y aumentaba el poder del empleador para expropiar el capital humano de los empleados específico de la empresa al amenazar implícitamente con despedirlos antes de que se crearan sus derechos de pensión, si insistían en un sueldo conmensurable con su valor para la compañía. En efecto, independientemente de las consideraciones del capital humano, podríamos imaginar a un empleador reduciendo el salario del empleado hasta el punto en que el salario y el beneficio de la pensión juntos excederían apenas al salario del empleado en su siguiente mejor empleo. El año anterior al retiro del empleado y su elegibilidad para la pensión, el salario podría ser cero… o aun negativo; el empleado pagaría para que se le permitiera seguir trabajando lo que fuese necesario para tener derecho a su pensión.

¿Significa esto que las prácticas de las pensiones eran explotadoras antes de ERISA? No necesariamente. Recuérdese que las condiciones del retiro, incluidos los derechos de pensión, son objeto de una negociación contractual entre el empleador y el empleado potencial, no una imposición unilateral.

[38] La Ley de Seguridad de los Ingresos de Jubilación del Empleado (ERISA) regula los beneficios de salud y bienestar de los empleados, así como las pensiones, y proporciona un foro federal para litigar disputas contractuales sobre los términos de las pensiones particulares u otros planes cubiertos.

Aun si el empleador se negara a negociar por separado con cada empleado, pero ofreciera condiciones de empleo sobre una base de tómalo o déjalo, la competencia entre los empleadores daría a los empleados potenciales una elección entre diferentes paquetes de salarios y prestaciones. Los paquetes ofrecidos por algunos empleadores destacarían los buenos beneficios del retiro o de otra clase a costa de los salarios menores, mientras que los ofrecidos por otros empleadores destacarían los salarios elevados a costa de beneficios de jubilación o de otra clase menos generosos o menos seguros. Los empleados se repartirían entre los empleadores de acuerdo con las preferencias de cada uno de ellos acerca del riesgo y la distribución del consumo a lo largo del ciclo vital.

Aun el otorgamiento incompleto no era un engaño ni un impedimento para la movilidad óptima de los trabajadores. Al hacer depender los beneficios de la pensión del hecho de que el empleado permaneciera en la empresa y se desempeñara satisfactoriamente, el otorgamiento incompleto facilitaba la recuperación por parte de los empleadores de su inversión en el capital humano de sus empleados específico de la empresa. Ello resolvía también el problema del último periodo del empleado mencionado en la sección anterior no sólo con el garrote (la amenaza del despido antes de que se crearan derechos de pensión) sino también con la zanahoria, ya que los beneficios de la pensión en un plan de beneficios definidos se ven fuertemente influidos por el salario del empleado en sus últimos años de trabajo. El incentivo del empleador para abusar del poder que le confería el otorgamiento incompleto, infringiendo su contrato no escrito de tratar justamente a sus empleados, se vería frenado por su deseo de preservar una reputación de trato justo (si perdiera eso tendría que pagar salarios mayores a sus empleados nuevos) y por el poder de negociación que confiere a un trabajador la posesión de capital humano específico de la empresa. En efecto, antes de ERISA eran raros los despidos oportunistas de trabajadores cubiertos por un plan de pensión, y esa ley no ha tenido ningún impacto detectable sobre los despidos de los trabajadores que cubre.[39]

Incluso antes de que se aboliera la jubilación obligatoria a una edad determinada, estaba lejos de ser universal. De hecho, se encontró principalmente en empresas que tenían planes de jubilación de beneficios definidos para sus empleados.

La jubilación obligatoria impondría altos costos a un trabajador que no tuviera un ingreso de jubilación seguro, lo que proporciona un plan de beneficios definidos. Por lo tanto, la combinación de la jubilación obligatoria con dicho plan protege los ingresos del trabajador, al mismo tiempo que permite

[39] Christopher Cornwell, Stuart Dorsey y Nasser Mehrzad, "Opportunistic Behavior by Firms in Implicit Pension Contracts", 26 *J. Human Resources,* 704 (1991); Richard A. Ippolito, "A Study of the Regulatory Impact of ERISA", 31 J. LAW & ECON., 85, 91-102 (1988).

al empleador despedir a los trabajadores mayores sin evaluar su productividad de forma individual.

ERISA requiere, entre otras cosas, que los planes de beneficios definidos creen los derechos de pensión del empleado después de que haya participado en el plan durante cinco años. Al limitar así el otorgamiento incompleto, la ley ha tendido a reducir el control de los empleadores sobre sus empleados más viejos. Se esperaría que tal pérdida de control llevara a los empleadores, primero, a invertir menos en el capital humano de sus empleados específico de la empresa y, segundo —en virtud de que los empleadores tendrían una inversión menor en los empleados que deberían proteger, y de que los empleados tendrían un incentivo menor para desempeñarse bien (ya que no afrontarían una pérdida sustancial de los beneficios de la pensión si fuesen despedidos)—, a recurrir con mayor frecuencia a una amenaza explícita o implícita del despido a fin de mantener la disciplina.

Se presentan situaciones difíciles en la aplicación de ERISA cuando un empleado que está inscrito en un plan de beneficios definidos deja a su empleador antes de la edad de retiro. Supongamos que tiene un saldo de efectivo en el plan de beneficios definidos (véase la nota 37 anterior). Por cada año que él permanezca en el plan, el monto de su cuenta nocional crecerá de acuerdo con la tasa de interés especificada en el plan. Cuando cumpla la edad de retiro él podría tener derecho a la cantidad total de su cuenta existente en ese momento, la cual podría retirar en su totalidad o en anualidades. Si abandona el empleo anticipadamente debiera tener derecho, como establece la ley, al valor presente de la cantidad que pudiese esperar tener en su cuenta cuando alcanzara la edad de retiro. Supongamos que desea la suma total en lugar de la anualidad. Calcular la suma total que sería el valor presente de lo que podría esperar en el retiro es algo capcioso o engañoso.[40] Esto requiere que incremente su balance de efectivo por la cantidad de interés que podría esperar recibir hasta el retiro y descontar después el incremento resultante del saldo de efectivo para reflejar el beneficio de recibir el dinero ahora en vez de en el futuro. A esta acción opuesta de incrementar el saldo de efectivo por el interés futuro sobre éste y descontar después ese aumento del saldo de efectivo al valor presente se le llama "serruchar". Esto involucra proyectar a futuro el valor del beneficio del retiro a la edad de retiro (ascendiendo) y después descontar de vuelta hacia el presente (descendiendo), y se asemeja así a la acción de dos personas con una sierra de dos mangos. Con frecuencia, el plan otorgará al participante el derecho a una tasa de interés variable antes que fija y, entonces, los intereses por crédito que se habrían acumulado entre la distribución de la suma global obtenida por un empleado al retirarse anticipadamente y los que obtiene al alcanzar la edad normal de jubilación no se

[40] Véase Ruppert *vs.* Alliant Energy Cash Balance Pension Plan, 2013 WL, 4034394 (7th Cir., 9 de agosto, 2013).

pueden conocer con certeza; sólo se pueden estimar. La tasa elegida para estimar esos intereses por crédito se llama "tasa de proyección". Ésta no generará una coincidencia exacta con los intereses por crédito porque las variaciones en los intereses por crédito futuros no se pueden conocer. Pero debe ser una estimación razonable, de buena fe. Sólo si la tasa de proyección que se utiliza para calcular el saldo de efectivo a la edad de retiro es idéntica a la tasa de interés que se utiliza para descontar el saldo de efectivo a valor presente (la tasa de descuento), podrá ser la suma global del jubilado anticipado simplemente el saldo de efectivo acumulado a la fecha en que deja el empleo del patrocinador del plan.

Lecturas recomendadas

Becker, Gary S., *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education* (3a ed., 1993).

Bird, Robert C. y John D. Knopf, "Do Wrongful-Discharge Laws Impair Firm Performance?", 52 *J. Law & Econ.*, 197 (2009).

Botero, Juan C., *et al.*, "The Regulation of Labor", 119 *Q. J. Econ.*, 1339 (2004).

Donohue III, John J. y Peter Siegelman, "The Changing Nature of Employment Discrimination Litigation", 43 *Stan L. Rev.,* 983 (1991).

Epstein, Richard A., "A Common Law for Labor Relations: A Critique of the New Deal Labor Legislation", 92 *Yale L. J.,* 1357 (1983).

———, "In Defense of the Contract at Will", 51 *U. Chi. L. R.,* 947 (1984).

Fischel, Daniel y John H. Langbein, "ERISA's Fundamental Contradiction: The Exclusive Benefit Rule", 55 *U. Chi. L. Rev.,* 1105 (1988).

Ippolito, Richard A., "The Implicit Pension Contract: Developments and New Directions", 22 *J. Human Resources,* 441 (1987).

Jolls, Christine, "Hands-Tying and the Age Discrimination in Employment Age", 74 *Tex. L. Rev.,* 1813 (1996).

———, "Antidiscrimination and Accommodation", 115 *Harv. L. Rev.,* 642 (2001).

Labor and Employment Law and Economics (Kenneth G. Dau-Schmidt, Seth D. Harris y Orly Lobel [comps.], 2009).

Lazear, Edward P., "Why Is There Mandatory Retirement?", 87 *J. Pol. Econ.,* 1261 (1979).

Lundberg, Shelly J., "The Enforcement of Equal Opportunity Law under Imperfect Information: Affirmative Action and Alternatives", 106 *Q. J. Econ.,* 309 (1991).

Posner, Richard A., *Aging and Old Age*, caps. 3, 13 (1995).

———, "An Economic Analysis of Sex Discrimination Laws", 56 *U. Chi. L. Rev.,* 1311 (1989).

Research Handbook on the Economics of Labor and Employment Law (Cynthia L. Estlund y Michael L. Wachter [comps.], 2012).

Schwab, Stewart J., "Life-Cycle Justice: Accommodating Just Cause and Employment at Will", 92 *Mich. L. Rev.,* 8 (1993).

Wachter, Michael L. y George M. Cohen, "The Law and Economics of Collective Bargaining: An Introduction and Application to the Problems of Subcontracting, Partial Closure, and Relocation", 136 *U. Pa. L. Rev.,* 1349 (1988).

Problemas

1. Se crea una unión sindical cerrada, proscrita por la Ley Taft-Hartley, cuando el empleador conviene con el sindicato en que él contratará sólo trabajadores que sean miembros del sindicato (lo que contrasta con una unión sindical, donde los trabajadores deben unirse al sindicato si, y cuando, se convierta en el representante de su unidad para la negociación colectiva). ¿Cuál es la función económica de la unión sindical cerrada?

2. ¿Es probable que los sindicatos sean más poderosos en las industrias de servicios que en las de manufacturas? ¿En el sector público o en el sector privado? Más generalmente, ¿cuáles son los factores que predisponen a una industria a la sindicalización?

3. ¿Qué tratan de maximizar exactamente los líderes de un sindicato? ¿El salario medio por hora de sus miembros? ¿El salario medio anual? ¿Los salarios totales de los trabajadores empleados en la actividad sindicalizada? ¿Cuáles efectos diferentes podrían tener estos diversos maximandos sobre el tamaño del sindicato? ¿Por qué podrían preocuparse los líderes del sindicato por el tamaño del sindicato y por los ingresos de sus miembros? ¿Cómo puede esto explicar la imposición de trabajadores innecesarios?

4. Los salarios han aumentado más de prisa en las ocupaciones no sindicalizadas, sobre todo en el servicio doméstico, que en las ocupaciones sindicalizadas. ¿Indica esto que la sindicalización es ineficaz para aumentar los salarios por arriba de los niveles competitivos?

5. ¿Pueden los sindicatos aumentar los salarios por arriba de los niveles competitivos a largo plazo si los empleadores venden en mercados de productos competitivos? ¿O si todos los trabajadores están sindicalizados?

6. Si un empleador fuese acusado de discriminar por la edad, ¿sería una defensa el hecho de que el demandante esté recibiendo un salario mayor que el de otros empleados de este empleador que estén haciendo el mismo trabajo?

7. Si un empleado logra probar que fue despedido en violación de una ley contra la discriminación, ¿cuál debiera ser la medida de los daños? ¿Debiera ser la diferencia existente entre su sueldo recibido de su empleador anterior y su sueldo actual (suponiendo que ha encontrado otro empleo)? ¿Bajo cuáles circunstancias compensaría en exceso esa medida al empleado?

XIII. EL SERVICIO PÚBLICO Y LA REGULACIÓN DEL MEDIO DE TRANSPORTE COLECTIVO[1]

XIII.1. El monopolio natural

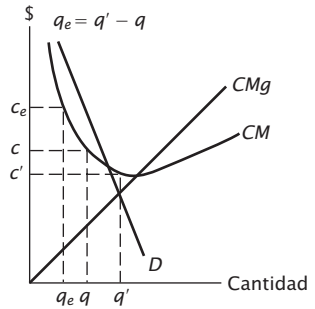
En la gráfica XIII.1 se describen las circunstancias estructurales en las que el monopolio es la forma más barata de organizar una industria (y por ello un monopolio *natural*). Obsérvese que el costo medio (*CM*) es declinante en el punto donde se interseca la curva del costo medio con la de la demanda. Supongamos que hay una empresa en el mercado que produce q unidades a un costo medio de c. Aumentaría la eficiencia llevando la producción hasta q', donde se igualarían la demanda y el costo marginal. Hay dos caminos para llegar a la producción mayor. Uno es que la empresa existente produzca $q' - q$ más unidades. Otro es que entre al mercado una empresa nueva y produzca q_e unidades (igual a $q' - q$). En q', el costo medio de la empresa existente sería c', pero el nuevo miembro tendría que incurrir en un costo medio de c_e para producir q_e unidades, y c_e se encuentra por encima de c'. Por lo tanto, resulta más barato que la empresa existente provea las unidades adicionales. Ello no se debe a que esa empresa sea más eficiente en el sentido de que su curva de costos se encuentre por debajo de las curvas de otras empresas —los cuadros de costos de la empresa existente y del recién ingresado son idénticos en la gráfica XIII.1—, sino porque una sola empresa puede proveer toda la producción demandada a un costo menor de lo que podrían hacerlo varias empresas.[2]

La razón es que los costos fijos son muy grandes en relación con la demanda. Si tales costos pueden repartirse entre toda la producción del mercado, una sola empresa que provea esa producción podría tener un costo medio de producción menor que dos empresas igualmente eficientes, cada una de las cuales incurriría en los mismos costos fijos, pero sería capaz de repartirlos entre sólo la mitad de la producción.[3] Esto es posible aunque (como en

[1] Véanse buenas introducciones al tema de este capítulo y el siguiente en Stephen G. Breyer, Richard B. Stewart y Cass R. Sunstein, *Administrative Law and Regulatory Policy* (6ª ed., 2004); Charles F. Phillips Jr., *The Regulation of Public Utilities: Theory and Practice* (3ª ed., 1993).

[2] Los costos promedio no tienen por qué estar reduciéndose realmente en la intersección con la curva de la demanda. Mientras dos o más empresas tengan costos medios mayores que una, el mercado es un monopolio natural; es decir, el mercado tiene "espacio" para una sola empresa.

[3] Por supuesto, una empresa grande puede ser menos eficiente que una pequeña a causa de la pérdida de control sobre los subordinados, la ausencia de vigor innovador u otras fuentes de las deseconomías de escala. Un monopolista natural es una empresa que tiene economías de escala *netas* en toda la gama de su producción viable. Ésta es una calificación importante, ya que plantea



GRÁFICA XIII.1

la gráfica XIII.1) el costo marginal esté aumentando con la producción. Para ilustrarlo, supongamos que los costos fijos de la producción de algún servicio sean 10 dólares, y que el costo marginal aumente como se indica en el cuadro XIII.1 (de donde se deriva la gráfica XIII.1). Si la producción del mercado es 6 o menos, una empresa podrá proveer esa producción a un costo total menor de lo que podrían hacerlo dos o más empresas (por ejemplo, a una producción de 6, el costo total si sólo hay una empresa es de 31 dólares, mientras que, si hubiera dos empresas, y cada una de ellas produjera tres unidades, el costo total sería de 32 dólares). Por lo tanto, la eficiencia requiere que sólo una empresa sirva a este mercado, a menos que la demanda sea de 7 o más unidades.

Un ejemplo plausible del monopolio natural es el servicio eléctrico local; la extensión del mercado está limitada debido al costo de transportación de la electricidad a largas distancias, mientras que los costos fijos (equipo de generación, una red de cables por toda la ciudad, etc.) son elevados. Pero, si un mercado es suficientemente pequeño, casi cualquier clase de empresa puede tener un monopolio natural —una tienda de abarrotes en una aldea, por ejemplo— porque todas las empresas tienen algunos costos fijos, los que podrían dominar a los costos totales si la demanda fuese suficientemente pequeña.

El monopolio natural plantea tres problemas que ameritan la regulación pública, según se ha pensado. Uno es la fijación de precios monopólicos. La empresa que abastece a un mercado de monopolio natural tiene el mismo incentivo que cualquier otra empresa para maximizar la ganancia restringiendo su producción, y tiene una probabilidad mucho mayor de alcanzar esta meta porque no necesita incurrir en los costos legales y administrativos de la colusión o de la adquisición de compañías para hacerlo. Las leyes antimonopólicas son ineficaces en esta situación; si el monopolista natural se

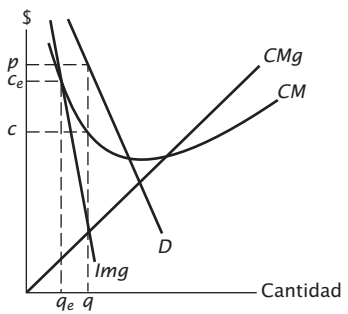
la cuestión de cuán importante es en efecto el monopolio natural en el terreno empírico: quizá no sea muy importante, a menos que, como se indica más adelante en el texto, la gama de la producción viable sea muy pequeña porque la demanda del mercado es muy pequeña.

CUADRO XIII.1

Producción	*Costos fijos*	*Costos totales fijos y variables*	*Costo marginal*	*Costo total medio*
1	$10	$11	$1	$11.0
2	10	13	2	6.5
3	10	16	3	5.3
4	10	20	4	5.0
5	10	25	5	5.0
6	10	31	6	5.2
7	10	38	7	5.4
8	10	46	8	5.8

divide, cada componente, con costos medios decrecientes, crecerá hasta que alguno tenga todo el mercado nuevamente.

El segundo problema (superficialmente inconsistente) es el del aliciente a la entrada ineficiente que se representa en la gráfica XIII.2. El monopolista vende la cantidad q al precio p, determinados por la intersección del costo marginal y el ingreso marginal. Un miembro recién ingresado, viendo que puede abastecer a una porción del mercado, q_e, a un costo medio (c_e) menor que el precio de mercado, tiene un incentivo para entrar. Cuando lo hace, la empresa existente deberá reducir su precio o disminuir su producción y, si sigue el último camino, el costo medio de la producción será mayor de lo necesario. Si por el contrario baja su precio, reducirá el precio del neófito, quien tendrá costos más altos debido a su escala de operaciones menor. Este precio menor aún podría ser remunerativo, aunque menor al costo del nuevo miembro, y, en ese caso, este último se verá obligado a salir del mercado y el monopolista será libre otra vez de cobrar el precio maximizador del benefi-



GRÁFICA XIII.2

cio. Eso puede sobrepasar su costo en gran medida, dependiendo de la elasticidad de la demanda de su producto.

El tercer problema es el de la dificultad para elaborar una estructura de precios eficiente. En el punto de las gráficas XIII.1 y XIII.2 donde el costo marginal se interseca con la demanda, el costo medio de la empresa es mayor que su costo marginal. Por lo tanto, si la empresa vendiera toda su producción al costo de la última unidad (costo marginal), su ingreso total sería menor a su costo total.[4] Pero si en consecuencia vende a un precio igual al costo medio, inducirá al comprador marginal a desplazarse a productos sustitutos, aunque su producción cueste más que llevar la producción del primer producto hasta el punto en el que el precio sea justamente igual al costo marginal. Como en cualquier caso de fijación de precios monopólicos, el comprador marginal se enfrentará a lo que, desde un punto de vista social, es una alternativa falsa.

La respuesta estadunidense tradicional al problema del monopolio natural era la regulación de los servicios públicos o del medio de transporte colectivo.[5] Esta regulación tiene tres elementos primarios: *1)* control de la ganancia (las tarifas de la empresa regulada no deberán exceder el nivel necesario para permitir que la empresa cubra su costo de servicio, incluido un rendimiento razonable sobre el capital invertido); *2)* control de la entrada (una empresa no podrá proveer un servicio regulado sin obtener primero, de la dependencia reguladora, un permiso), y *3)* control de la estructura de precios (la empresa no podrá discriminar en sus tarifas). Los tres controles se refieren a los tres problemas del monopolio natural esbozados antes. Sin embargo, muchos de los servicios que habían estado sujetos a tal regulación no eran monopolios naturales y podrían existir mejores métodos para ocuparse de los que sí lo son. A partir del decenio de 1970, el movimiento de desregulación redujo o eliminó la regulación de los servicios públicos y los medios de transporte colectivo en la transportación aérea y terrestre, en los servicios de telecomunicaciones, en los financieros y en los energéticos. El movimiento de desregulación se aborda en este capítulo y en el siguiente.

XIII.2. El control de las ganancias y el problema del rendimiento razonable

En el método tradicional de limitar las ganancias de las empresas reguladas, la dependencia reguladora selecciona primero un año reciente representativo de

[4] El costo total es igual al costo promedio multiplicado por la cantidad vendida. Si el costo marginal es menor que el costo promedio, el producto del costo marginal multiplicado por la cantidad vendida será menor que el costo total.

[5] El término de "medio de transporte colectivo" se emplea de ordinario para una empresa de transporte de pasajeros o de comunicaciones, mientras que el de "servicio público" se asigna a un productor o proveedor de electricidad o de gas, pero los regímenes reguladores son de ordinario similares.

las operaciones de la compañía y calcula los costos totales en que ha incurrido la empresa para proveer en ese año el servicio regulado, menos todos los costos de capital, incluyendo los intereses de la deuda a largo plazo. La suma de los costos incluidos se designa como el costo del servicio de la compañía en el año de prueba. La dependencia añade al costo del servicio un rendimiento razonable sobre el capital invertido (la deuda a largo plazo más el capital social). El componente del rendimiento se determina multiplicando la base tarifaria de la compañía —una estimación del valor de los activos de capital empleados por la compañía para prestar el servicio regulado— por la tasa de rendimiento razonable de la compañía. Esa tasa es a su vez un promedio ponderado de la tasa de interés a largo plazo más la tasa de rendimiento para los accionistas del capital social que la dependencia considere apropiada en vista del riesgo de la inversión y de la tasa de rendimiento disfrutada por los accionistas de empresas comparables. El componente del rendimiento, sumado al costo del servicio, establece los requerimientos de ingresos de la compañía. La compañía presenta a la dependencia una lista de tarifas destinada a generar justamente esta cantidad de ingresos, bajo el supuesto de que la cantidad demandada del servicio de la empresa será la misma que en el año de prueba. Una vez aprobadas, estas tarifas son el máximo que la empresa puede cobrar[6] y también son, a menudo, mínimas. Véase después XIII.5.

La determinación del costo del servicio en el año de prueba es relativamente simple, aunque algunas ganancias monopólicas pueden ocultarse en rubros del gasto tales como sueldos y las prestaciones de los administradores.[7] Pero la determinación del componente del rendimiento plantea dificultades considerables. La cuestión que preocupó durante largo tiempo a los tribunales y a los jurisconsultos era si la base tarifaria debería medirse por el costo original depreciado de los activos de capital de la empresa o por el de de reposición. La cuestión es importante en los periodos de inflación prolongada, cuando el costo de reposición de un activo de capital de larga vida puede exceder grandemente al costo original del activo, de modo que el empleo del costo de reposición facultaría a la compañía para recibir una tasa de rendimiento mucho mayor, lo que a su vez se traduciría en tarifas más altas. Supongamos que una empresa paga 10 000 dólares por una máquina que puede producir 1 000 unidades por año y tiene una vida útil de 20 años. Los costos de operación son de 1 dólar por unidad y el costo anual del capital usado para comprar la máquina es de 5%. Por lo tanto, el costo medio es

[6] Hay dos procedimientos diferentes para cambiar las tarifas. Primero, la dependencia puede iniciar un procedimiento como el descrito en el texto, para revisar las tarifas de la empresa. Segundo, y más comúnmente, la empresa puede iniciar un procedimiento para que se le dé permiso de aumentar las tarifas. Éste es de ordinario un procedimiento más abreviado. La carga de la justificación de un aumento recae sobre la empresa.

[7] Véanse después XIII.3 y XXVIII.1.

de 2 dólares.[8] Después de 10 años aumenta la demanda del producto de la empresa, de modo que decide comprar una segunda máquina. Pero los precios han aumentado. Una máquina idéntica cuesta ahora 15 000 dólares y, mientras que los costos de operación permanecen constantes, el costo de capital ha aumentado en 6%. Por lo tanto, el costo medio de la producción con la segunda máquina es 2.65 dólares. ¿Cuál debería ser el precio de la empresa: 2.00, 2.325 o 2.65 dólares?

El último. Un precio menor puede inducir a la empresa a tomar una decisión de inversión incorrecta. Supongamos que el precio es de 2.33 dólares. Quienes valúen el producto en más de 2.33 pero menos de 2.65 dólares tratarán de comprarlo. Frente a una demanda excesiva, la empresa podría comprar una tercera máquina. Cuando aumente entonces su precio para cubrir el costo de producción de esa máquina (el que será también de 2.65 dólares, podemos suponer), perderá a muchos de sus clientes nuevos; se habrá expandido en exceso.

Si el costo de reposición (es decir, el corriente), antes que el costo original, es estándar económico correcto, ¿por qué ha sido controversial la cuestión del costo original contra el costo de reposición? Hay tres razones para ello. Primero, el costo de reposición no es siempre el estándar económico correcto. Si los principales activos de capital de una industria (por ejemplo, las vías ferroviarias, las locomotoras y las terminales) no van a ser completamente remplazadas, debido a la declinación de la demanda del producto de la industria, el cobro de un precio basado en los costos de reposición sólo impedirá la utilización plena de la capacidad existente al alejar a los consumidores; porque no tiene caso confrontar a los consumidores con el costo de reposición de los activos que no se van a remplazar. El estándar apropiado en tal caso es el costo de oportunidad de los activos existentes en la industria; si, como en el caso del ferrocarril, los activos son muy especializados (no pueden trasladarse fácilmente a otro uso), ese costo podría ser mucho menor que el costo de su reposición (véase después XIII.3).

Segundo, cuando el costo original es menor que el costo de reposición, el uso de la última medida para determinar los requerimientos de ingresos de la empresa hará que los ingresos de ésta excedan a los costos anotados en los libros de la empresa, lo que producirá una ganancia súbita aparente para los accionistas de dicha empresa. Pero la ganancia súbita podría ser una ilusión creada por la inflación.[9] Y en todo caso se ve contrarrestada por las pérdidas de los accionistas y acreedores de empresas de otras industrias, tales como los

[8] La depreciación es 10 000 dólares ÷ 20, o sea, 500 dólares por año, que es igual a 50 centavos por unidad. El costo anual del capital, 5% de 10 000 dólares, es de otros 50 centavos por unidad, y los costos de operación son de 1 dólar por unidad.

[9] O bien podría derivar de un cambio en las condiciones subyacentes de la oferta o la demanda, sin relación con la inflación, que ha provocado un aumento en el precio real del bien. En tal caso, la "ganancia súbita" es una forma de la renta económica. Véase antes I.1.

ferrocarriles, donde el valor de los recursos especializados se ha derrumbado porque la demanda del producto de la industria ha disminuido en lugar de aumentar. *Ex ante*, no hay ninguna ganancia súbita.

Tercero, la cuestión legal de los costos corrientes contra los costos de reposición surge en una forma que oscurece la realidad económica. Se ha sostenido que la garantía constitucional de la compensación justa por las expropiaciones gubernamentales faculta a los accionistas de un servicio público para recibir un rendimiento "justo" de su inversión. Dos consideraciones son relevantes al decidir si la utilización del costo original para la determinación de la tarifa máxima de un servicio público impide que los accionistas obtengan un rendimiento justo. La primera consiste en saber si los accionistas conocían el estándar que se aplicaría cuando invirtieron. Si sabían que la dependencia reguladora aplicaba un estándar de costo original, no podrán alegar que hubo una confiscación, pues presumiblemente, aun con esta circunstancia, su inversión en la empresa regulada era más atractiva que sus oportunidades de inversión alternativas, o de otro modo no habrían realizado la inversión.[10] Si invirtieron bajo el supuesto razonable de que la dependencia aplicaría un estándar de costo de reposición, la acusación de confiscación resultará más plausible.

La segunda consideración consiste en saber si los rendimientos adicionales para los accionistas, resultantes del empleo de un estándar de costo de reposición, se encuentran dentro del alcance de la garantía constitucional de la compensación justa. Esto podría depender de que tales rendimientos se caractericen o no como ganancias súbitas. Ambas consideraciones desvían la atención de la cuestión de la sensatez económica del estándar del costo original.

Otro problema importante de la fijación del componente del rendimiento en los requerimientos de ingresos de la empresa regulada es el de la medición del costo del capital social. Pero pospondremos la consideración de este problema hasta el capítulo que se ocupa de la teoría del mercado de capital (véase después XVI.7).

XIII.3. Otros problemas causados por el intento de limitar las ganancias de una empresa regulada

El empleo de los costos pasados, antes que de los corrientes como base para la fijación de precios de la empresa regulada, resulta muy problemático. Si los costos han aumentado desde el año de prueba, las tarifas serán demasiado

[10] Si los inversionistas son tenedores de bonos, la empresa regulada está pagándoles presumiblemente una tasa de interés suficientemente elevada para compensarlos por las desventajas que pueda imponerles el estándar del costo original. Si son accionistas, la tasa de rendimiento sobre el capital social que se esperaba que fuese permitida por la dependencia era presumiblemente, cuando ellos invirtieron, lo bastante elevada para que la inversión fuese rentable a pesar de que la dependencia empleara el estándar del costo original.

bajas para cubrir los costos corrientes de la empresa, y es posible que la empresa no pueda recuperarse solicitando aumentos de las tarifas. Si los costos han bajado, las tarifas corrientes serán demasiado elevadas. Si la demanda ha aumentado, las tarifas pueden ser de nuevo demasiado elevadas, ya que se calculan dividiendo los requerimientos de ingresos por la producción del año de prueba, y el costo medio de una producción mayor será menor si la empresa está operando en una región de costos medios declinantes. Adviértase también que la brecha introducida por el empleo de un año de prueba genera oportunidades para que la empresa engañe: reduciendo la calidad de su producción, puede reducir sus costos corrientes por debajo de sus costos del año de prueba, y aumentar en consecuencia sus ganancias.

La popularidad de los conceptos del costo original y el año de prueba del servicio es una manifestación de la lucha eterna, en la regulación, entre la teoría económica y la viabilidad. Ambos métodos facilitan el empleo de los propios libros contables de la compañía como la fuente de información básica para el proceso de elaboración de las tarifas, pero ambos conducen a graves alejamientos de la determinación eficiente de los precios.

Cuando la inflación es ligera y los precios reales de los insumos son constantes o decrecientes, la regulación de los servicios públicos parece tener escaso efecto sobre los niveles de los precios o de las ganancias, lo que implica que es un fracaso como sistema para el control del monopolio. Cuando la inflación es aguda, cuando la demora reguladora opera dramáticamente en contra de la empresa, las ganancias pueden disminuir por debajo del nivel competitivo, lo que implica una corrección excesiva. Y los errores implicados en los esfuerzos reguladores para limitar los precios y las ganancias no se cancelan recíprocamente. Si las tarifas son demasiado elevadas, los consumidores se verán inducidos a comprar productos sustitutos cuya producción es efectivamente más costosa para la sociedad. Si las tarifas son demasiado bajas, los consumidores se alejarán de los productos sustitutos cuya producción es menor para la sociedad.

En la medida en que la regulación de los servicios públicos alcance su objetivo de determinar el precio del servicio público sobre la base de cierta tasa de ganancia sobre el costo, reducirá el incentivo de la empresa para minimizar los costos (véase más sobre este punto, más adelante, en XV.17). El castigo para el monopolista que no pueda minimizar sus costos es que sus ganancias serán menores que de otro modo, y la regulación del servicio público tiende a reducir ese castigo. Sin embargo, el retraso regulador podría frenar esta tendencia, ¡lo que demuestra que la demora burocrática puede ser afortunada!

Dado que es más fácil para la dependencia monitorear los sueldos que las prestaciones, los administradores de una empresa regulada podrían sustituir los primeros con las últimas. La sustitución reduce el valor si un dólar gastado en prestaciones vale menos para los receptores que el mismo dólar que

se les entregue en efectivo. La sustitución es aún más atractiva por el hecho de que los beneficios marginales no son gravables para los beneficiarios.

De igual modo, si el costo del capital social no se encuentra tan eficazmente constreñido como los costos laborales, la empresa regulada tendrá un incentivo para emplear demasiado capital en relación con la mano de obra.[11] El incentivo podría explicar una leyenda famosa de los ferrocarriles: la renuencia de un ferrocarril a "recortarse" a sí mismo.[12] Supongamos que el ferrocarril X tiene una línea del punto A al punto C, una distancia de unos 1 290 kilómetros, y una línea de A a B, una distancia de unos 640 kilómetros; pero el ferrocarril Y tiene una línea de B a C, una distancia de sólo aproximadamente 320 kilómetros, de modo que resulta más rápido y barato ir de A a C a través de B, proveyendo los ferrocarriles X y Y un servicio conjunto, que utilizar la línea única más larga del ferrocarril X. Incluso si el ferrocarril X es el único que sirve a A y controla la elección de la ruta, escogerá la ruta más eficiente A-B-C (en ausencia de regulación), aunque esto signifique "recortarse" a sí mismo. X puede captar todos los rendimientos monopólicos derivados del tráfico, cualquiera que sea la ruta que se escoja, y estos rendimientos serán mayores cuando ofrezca el servicio conjunto con Y, porque la distancia más corta reduce los costos totales del servicio y aumenta la rapidez (por ende el valor) del servicio prestado a los clientes.[13] Pero, con la regulación de la tarifa, el recorte reducirá los costos de capital del ferrocarril y, en consecuencia, no podrá justificar una tarifa tan alta por el servicio. Por lo tanto, si los costos de capital están regulados imperfectamente, el ferrocarril podría preferir la ruta menos eficiente, que utiliza más de su propio capital.

Paradójicamente, también podría crearse un incentivo para invertir en exceso si la dependencia reguladora deprime las tarifas y las ganancias de la empresa regulada insistiendo en que ésta utilice el costo original depreciado, antes que el costo de reposición, para determinar la cantidad por la que se multiplicará la tasa de rendimiento permitida a fin de obtener las ganancias autorizadas a la empresa. La empresa tendrá un incentivo para expandir su planta a fin de servir a nuevos clientes, e incluir el costo de la nueva inversión en su tarifa básica. Los clientes antiguos pagarán más porque el costo de capital de la empresa es mayor, pero el aumento de la tarifa será relativamente indoloro porque estaban pagando tarifas bajas a resultas de la política de la dependencia reguladora para la valuación de la tarifa básica. Los clientes nuevos podrían pagar menos que el costo incremental de la expansión porque

[11] Harvey Averch y Leland L. Johnson, "Behavior of the Firm under Regulatory Constraint", 52 *Am. Econ. Rev.*, 1052 (1962). Pero véase Paul L. Joskow y Roger G. Noll, "Regulation in Theory and Practice: An Overview", en *Studies in Public Regulation* 1, 10-14 (Gary Fromm [comp.], 1981).

[12] Véase, por ejemplo, Missouri-Kansas-Texas R. R. *vs.* United States, 632 F.2d, 392, 404-406 (5th Cir., 1980).

[13] Sin embargo, ésta es una situación clásica de monopolio bilateral, y los altos costos de transacción en tal situación podrían hacer que X prefiriera la ruta "menos eficiente" (haciendo a un lado los costos de transacción), incluso en ausencia de regulación. Véase III.7 anterior.

ese costo se está promediando entre todos los clientes de la empresa; pero a la empresa no le importará porque estará recuperando los costos plenos de la expansión de manos de sus clientes en conjunto. El resultado será una construcción excesiva.[14]

Una dependencia reguladora tendrá menos capacidad para controlar las ganancias de una empresa si ésta tiene filiales no reguladas, a las que la empresa podría ser capaz de asignar algunas de las ganancias de su servicio regulado; por lo tanto, la regulación crea incentivos para la expansión a otros mercados aun cuando tal expansión sea ineficiente. Esto podría impedirse prohibiendo que una empresa regulada opere en mercados no regulados, pero tal prohibición podría impedir la integración eficiente.

La regulación de las ganancias podría inducir a la empresa regulada a inclinar su estructura de capital hacia la deuda; aun cuando esto aumentará el riesgo de quiebra de la empresa, porque el servicio de la deuda es un costo fijo y, por lo tanto, no puede evitarse mediante la reducción de la producción de la empresa si la demanda disminuye, el riesgo bien puede valer la pena. La empresa sabe que la dependencia reguladora se sentiría avergonzada por el hecho de que se fuera a la quiebra porque es probable que la dependencia fuese culpada por no haber permitido que la empresa cobrara tarifas remunerativas; y, por lo tanto, al aumentar el riesgo de quiebra de la empresa, la estructura de capital riesgoso podría disuadir a la dependencia de obligar a la empresa a reducir sus tarifas.[15]

¿Qué diremos si un proyecto de capital de una empresa de servicio público se cancela inesperadamente, como cuando los reguladores obligan al cierre de una planta de energía nuclear parcialmente terminada debido a preocupaciones por la seguridad?[16] ¿Debiera permitirse a la empresa que incluyera en la tarifa básica su inversión en la planta (supongamos que tal inversión era prudente cuando se realizó), trasladando así a sus clientes el costo del cierre? La respuesta debiera depender de la medida en que la regulación del servicio público alcance su objetivo formal de determinar los precios sobre la base de cierta ganancia por encima del costo. Si lo logra plenamente, los clientes obtendrán el beneficio de cualesquiera disminuciones inesperadas de los costos y, por lo tanto, debieran soportar la carga de cualesquiera aumentos inesperados de los costos, siempre que la empresa esté actuando prudentemente. De otro modo, los consumidores soportan todas las ganancias inesperadas y los accionistas todas las pérdidas inesperadas, y eso desembocará en una inversión subóptima en el servicio regulado.

[14] Véase Daniel F. Spulber, "Pricing and the Incentive to Invest in Pipelines after *Great Lakes*", 15 *Energy L. J.*, 377 (1994).

[15] Véanse pruebas de este fenómeno en Yossef Spiegel y Daniel F. Spulber, "The Capital Structure of a Regulated Firm", 25 *RAND J. Econ.,* 424 (1994).

[16] Véase Martin B. Zimmerman, "Regulatory Treatment of Abandoned Property: Incentive Effects and Policy Issues", 31 *J. Law & Econ.,* 127 (1988).

Ya debe ser evidente que el problema básico de la regulación es el de la propiedad dividida. Los derechos de propiedad de los activos de la empresa regulada están en efecto divididos entre la empresa regulada y la dependencia reguladora. Dado que el problema de la propiedad dividida es recurrente en las áreas de la conducta reguladas por el derecho común, es posible que haya una alternativa del derecho común a la regulación de las empresas de servicios públicos. En el capítulo siguiente exploraremos esta posibilidad (véase después XIV.11).

XIII.4. La regulación de los incentivos

La demora de la regulación es un método imperfecto para recompensar al monopolista regulado eficiente y castigar al ineficiente. Los costos de la empresa regulada podrían estar aumentando por razones no relacionadas con su eficiencia, tales como un aumento de precios en un insumo fundamental, o disminuyendo por razones no relacionadas con su eficiencia, tales como un aumento de la demanda que le permita repartir sus costos fijos sobre una producción más abundante. La demora de la regulación confiere pérdidas súbitas en el primer caso y ganancias súbitas en el segundo.

Un procedimiento para ligar más estrechamente el nivel de ganancias permitido al desempeño de la empresa consiste en permitir cambios de la tarifa basados en los cambios ocurridos en un índice de costos de toda la industria. Supongamos que 20 empresas de servicio público sirven a áreas diferentes, pero son comparables en sus características de tamaño y operación, y que en un periodo el costo medio por kilovatio de estas empresas aumenta en un centavo de dólar; bajo el enfoque del índice, a cada una de estas empresas se le permitirá que aumente sus tarifas en un centavo (y no más) durante el periodo siguiente. Las empresas cuyo desempeño se encuentre por arriba del promedio en el periodo anterior obtendrán así una ganancia adicional; quienes se desempeñen por debajo del promedio serán castigadas.

Este proceso simula el de la competencia en el mercado, donde las empresas que tienen costos menores obtienen ganancias por encima del nivel competitivo a corto plazo. A largo plazo, a medida que las empresas más eficientes expanden su producción para cerrar la brecha existente entre el costo marginal y el precio del mercado, el nivel de precios baja y, a fin de continuar obteniendo ganancias por encima del nivel competitivo, las empresas tendrán que encontrar nuevos procedimientos para reducir sus costos. De igual modo, bajo el enfoque del índice (o "regulación del precio tope"), los esfuerzos de las empresas por superar al promedio con mejores controles de costos hacen que baje el promedio (tanto directamente como induciendo a las empresas que se encuentran de modo persistente por debajo del promedio a que cambien a

sus administradores), creando así incesantes presiones para la continuación de las reducciones de costos.[17]

Debe distinguirse este enfoque del de los traslados automáticos de los costos de insumos de combustibles que se volvieron populares cuando se dispararon los precios de los combustibles en el decenio de 1970. El enfoque de los traslados permite que la empresa regulada añada un dólar a sus tarifas (sin utilizar un procedimiento tarifario formal) siempre que un aumento de los precios de los combustibles añada un dólar a sus costos del servicio. Esto redujo el incentivo para economizar en el uso del combustible, negociar un precio mejor, cambiar a un combustible más barato o tomar otras medidas economizadoras.

En la práctica, la regulación de incentivos no es un sustituto perfecto del mercado. Un problema particularmente grave es el incentivo de una empresa con precio limitado para reducir la calidad a fin de reducir sus costos sin reducir el precio.[18] Éste es un problema general con las medidas basadas en el desempeño: un empleado al que se le paga una tasa fija por unidad de producción tendrá un incentivo para sustituir la calidad por la cantidad porque puede aumentar su producción al reducir aquélla.

Además, la regulación de incentivos no puede tener en cuenta, sin volver a la regulación del costo de servicio, los factores idiosincrásicos que aumentan los costos de una empresa regulada. En un mercado, una empresa que experimenta un costo no asumido por sus competidores está limitada, pero no necesariamente a su precio actual. Si hay diferenciación de producto, de manera que los productos de sus competidores no sean sustitutos perfectos para los suyos, puede cobrar al menos un precio ligeramente más alto al reducir la producción (es decir, enfrenta una curva de demanda de pendiente negativa, aunque sólo se inclina suavemente hacia abajo). Una empresa excluida por un tope de precios de hacer tal ajuste podría enfrentar la bancarrota, un destino que las dependencias reguladoras son reacias a reconocer; y, por lo tanto, hay una gran presión para conceder excepciones al tope de precios.

XIII.5. Regulación de la estructura tarifaria y de la entrada

A primera vista parece extraña la conjunción del control de las ganancias y la regulación de la entrada, ya que, si la fijación de precios monopólicos se

[17] La regulación del precio tope es el método empleado ahora para determinar la porción de los costos hospitalarios que será rembolsada por el gobierno federal en los términos de los programas de Medicare y Medicaid. La regulación del precio tope y otros métodos de regulación de los incentivos se describen en *Incentive Regulation for Public Utilities* (Michael A. Crew [comp.], 1994).

[18] Véase Noel D. Uri, "Measuring the Change in Productive Efficiency in Telecommunications in the USA", 29 *J. Econ. Stud.*, 150 (2002); Uri, "Service Quality Effects of Incentive Regulation on Access Service in Telecommunications in the United States", 16 *Euro. J. Law & Econ.*, 369 (2003).

evita a través de la regulación tarifaria, el nuevo miembro puede ganarse un sitio en el mercado del monopolista sólo cuando sus costos son menores que los del monopolista. O la regulación de la entrada se basa en una conciencia refinada de que el control del nivel general de los precios de los monopolistas regulados es a menudo ineficaz o, lo que es más probable, tiene propósitos completamente diferentes. Uno de tales propósitos podría ser el apoyo al tercer eslabón principal de la regulación de las empresas de servicios públicos y el medio de transporte colectivo, que es el control de la estructura de precios: cómo se convierten en precios específicos los requerimientos de ingresos de la empresa.

No hay ninguna respuesta satisfactoria para la interrogante, planteada antes, del precio óptimo de los servicios cuando el costo marginal es menor que el costo medio.[19] Igualar el precio al costo medio es poco satisfactorio por las razones que discutimos antes. Anteriormente se pensaba que la mejor solución era vender el servicio al costo marginal (a corto plazo), pagando el gobierno la diferencia resultante de la incapacidad de la empresa para recuperar su costo total con las recaudaciones fiscales generales. Pero esta solución tiene dos desventajas graves. Primero, al elevar las tasas fiscales en el resto de la economía, causa la misma clase de distorsiones de la asignación que, se supone, trata de curar (véase después XIII.7). Segundo, alienta a los consumidores para que utilicen los servicios producidos en condiciones de costo medio declinante en lugar de los servicios producidos en condiciones de costo medio creciente, aun cuando sea más costosa la producción de aquéllos.

Supongamos que no cuesta más construir y operar un puente que pueda soportar 1 000 automóviles por día que un puente que pueda soportar 900. La demanda diaria estimada, dentro de una gama amplia de precios posibles, es de 950 automóviles, más o menos. Los costos anuales de depreciación y operación del puente ascienden a un millón de dólares. Mientras se soporten menos de 1 000 automóviles, el costo marginal (es decir, el costo de aumentar la producción del puente en un viaje automovilístico más) será igual a cero. Pero supongamos que puede soportarse el mismo número de automóviles a la misma velocidad en una barcaza o transbordador con un costo total anual de sólo 100 000 dólares, pero que el costo marginal sería igual a 5 dólares. No habría ninguna demanda del servicio de la barcaza o transbordador siempre que el gobierno construyera o pagara el puente y cobrara una tarifa igual al costo marginal, es decir, que no cobrara nada. Por supuesto, si el gobierno fuese omnisciente y pudiera resistir las presiones de los consumidores, no construiría el puente porque sabría que existía un método socialmente menos costoso para proporcionar el mismo servicio. Pero el gobierno no es omnisciente. Y, si el puente se hubiese construido en una época en que el servi-

[19] Véanse discusiones aclaradoras del problema en Ronald H. Coase, "The Marginal Cost Controversy", 13 *Economica* 169 (1946); Coase, "The Theory of Public Utility Pricing and Its Application", 1 *Bell J. Econ. & Mgmt. Sci.,* 113 (1970).

cio de la barcaza o transbordador no era viable, y surgiera luego la interrogante de si conviniera reponer el puente, el gobierno se vería fuertemente presionado por quienes se hubiesen adaptado al servicio gratuito del puente para que fuese repuesto. El congestionamiento podría ser un problema si el precio del servicio del puente continuara siendo cero. Pero, incluso si el gobierno estuviera comprometido a fijar precios de cero cuando el costo marginal fuera cero, el costo marginal sería positivo si hubiera congestionamiento, ya que los costos de éste dependerían del número de personas que usan el puente (y cuándo, etcétera).

Una alternativa para igualar el precio al costo marginal es el precio en dos partes. Podría requerirse que el usuario del puente pagara: *1)* una cuota inicial única, calculada de tal modo que la suma de tales cuotas sufragara los costos fijos del puente, y *2)* una tarifa, igual al costo marginal, pagadera cada vez que cruzara el puente. Este método de fijación del precio permitiría que se cubrieran los costos fijos del puente y que se hiciera la compra marginal. Pero no permitiría que el comprador marginal —el comprador que estuviese dispuesto a pagar el costo marginal de su uso del puente, pero no a contribuir con su parte proporcional de los costos fijos— obtuviera el servicio. Supongamos que el cobro único es de 10 dólares y la tarifa (igual al costo marginal) es de 2 centavos de dólar (no cero, como en nuestro ejemplo anterior). Nuestro comprador marginal podría estar dispuesto a pagar 1 dólar por una sola vez pero no más, porque no espera usar el puente a menudo. Si se le niega el servicio, la compañía dueña del puente se encontrará en peor situación porque el comprador marginal habría hecho *alguna* contribución a sus costos fijos, y él mismo estaría también en peor situación, por supuesto.

Otra alternativa es el precio de Ramsey, que en su forma original significaba el cobro de un precio mayor entre menos elástica fuese la demanda del comprador. Al revés de lo que ocurre con el precio de dos partes, donde el precio medio pagado por cada comprador disminuye entre más compre, esta forma del precio de Ramsey es uniforme para cada comprador pero difiere entre los compradores, porque los demandantes menos elásticos pagarán más y los más elásticos pagarán menos. La mejor forma del precio de Ramsey —si omitimos los formidables costos de información implicados por los esfuerzos para medir las elasticidades e impedir el arbitraje— es la siguiente: como en el precio de dos partes, cada comprador paga una cuota de entrada para cubrir los costos fijos, pero la cuota varía inversamente con la elasticidad de la demanda del comprador… y el comprador verdaderamente marginal no paga ninguna cuota de entrada. Además de la cuota de entrada, cada comprador paga el costo marginal de cada unidad que compre. Dada una información perfecta, la producción se llevará hasta el punto en que el costo marginal se interseque con la demanda, pero sin imponer un déficit a la empresa regulada o un impuesto a quienes no usen el servicio regulado.

El precio en dos partes y el precio de Ramsey se asemejan pero no son idénticos a la discriminación de precios ordinaria; es decir, la discriminación

de precios que realiza un monopolista que maximiza sus ganancias (véase antes IX.5). (El precio óptimo de Ramsey, que acabamos de describir, se asemeja a la discriminación perfecta de los precios.) También bajo la discriminación de precios ordinaria varía el precio con la disposición a pagar, y en la misma dirección (es decir, el precio es mayor entre menos elástica sea la demanda del cliente). La semejanza resulta insólita a veces, como cuando el monopolista de un mimeógrafo exige que los usuarios de su máquina le compren la tinta. El precio de compra o de arrendamiento de la máquina corresponde al cobro fijo en un sistema de precios de dos partes; el precio de la tinta es el cobro variable, y es mayor entre más use el cliente la máquina. Sin embargo, en virtud de que el propósito de la discriminación de precios ordinaria no es el de permitir que se recuperen los costos fijos de una manera que permita el servicio a los compradores marginales, sino el de maximizar el exceso de los ingresos sobre los costos, el precio medio será mayor que bajo el sistema de precios de dos partes, y es probable que la producción sea menor porque el componente fijo del precio de dos partes podría disuadir a muchos demandantes.

Pero, al igual que la discriminación de precios ordinaria, el precio (de Ramsey) de dos partes puede distorsionar la competencia entre los compradores. Aun cuando se fije a cada comprador la misma porción de los costos fijos, el precio medio variará entre los compradores según la cantidad comprada (entre mayor sea la cantidad, menor será el precio medio), aunque no hay ninguna diferencia en el costo del servicio a compradores diferentes. Si están en competencia, algunos de ellos tendrán una ventaja competitiva no relacionada con la eficiencia superior.

Mucho tiempo antes de que hubiera elasticidades medidas de la demanda, la determinación de los precios de los servicios públicos y del medio de transporte colectivo se aproximaba a los precios de Ramsey. Un ejemplo es el de la fijación de las tarifas por el valor del servicio en la industria ferroviaria antes de la aparición del camión.[20] Mientras que las tarifas ferroviarias basadas en el costo medio serían aproximadamente proporcionales al *peso* de los embarques, las tarifas ferroviarias basadas en los principios de Ramsey serían, en ciertas circunstancias, aproximadamente proporcionales al *valor* de los embarques. Supongamos que el costo marginal de la transportación de cada uno de dos bienes por ferrocarril a la misma distancia en el mismo tiempo sea 1 dólar y que los ferrocarriles tratan de imaginar cómo cubrirán 3 dólares en costos fijos al mismo tiempo que se maximiza la producción. El valor de un embarque (de carbón) es de 10 dólares y el de otro (de cobre) es de 100. Si el ferrocarril cobra a cada cliente el mismo precio, 2.50 dólares (asignando así la mitad de los costos fijos a cada uno), el embarcador de carbón percibirá esto como el equivalente a un impuesto de 25% sobre su embarque y buscará

[20] Sylvester Damus, "Two-Part Tariffs and Optimum Taxation: The Case of Railway Rates", 71 *Am. Econ. Rev.,* 65 (1981).

otros modos de transportación. Si el ferrocarril cobra al embarcador de carbón sólo 1 dólar y al embarcador de cobre 4 dólares, el impuesto para el primero bajará a 10% pero el impuesto para el último será de sólo 4 por ciento.

Este esquema de fijación de precios sólo funciona si el embarcador de cobre no tiene buenas alternativas de transportación, como era cierto sólo en los primeros días de los ferrocarriles, antes de los camiones. Los embarcadores que se benefician de la fijación de los precios por el valor del servicio —los embarcadores de bienes pesados, de bajo valor— se resistirán naturalmente a un reajuste de las tarifas para reflejar las elasticidades cambiantes de la demanda de transportación ferroviaria, como lo hacen los modos de transportación rivales. Al ser órganos políticos, las comisiones reguladoras suelen rendirse a tales presiones. La decisión de la Corte Suprema sobre los moldes de lingotes[21] confirmó un fallo de la Comisión de Comercio Interestatal que prohibía a los ferrocarriles bajar el precio más allá de su costo medio a fin de arrebatar clientes a las líneas de lanchones. A primera vista, la competencia entre el ferrocarril y el lanchón se parece mucho a nuestro ejemplo del puente y la barcaza o transbordador. Los ferrocarriles, como los puentes, tienen grandes costos fijos (los derechos de vía, las vías, el equipo rodante, etc.) y costos marginales bajos; los lanchones, como las barcazas o transbordadores, tienen costos fijos bajos y costos marginales elevados. Pero éste no es un argumento a favor de que se exija a los ferrocarriles igualar los precios a los costos medios (a menos que se considere preeminente la preocupación por la distorsión competitiva en los mercados de compradores). Es un argumento a favor de que se exija a los ferrocarriles el uso de la regla de Ramsey antes que la igualación del precio al costo marginal, lo que genera tarifas bajas en sus mercados competitivos, donde, como vimos en el capítulo x, la elasticidad firme de la demanda de la empresa será elevada.

Hay un argumento a favor de que se permita a los ferrocarriles cobrar precios aún menores: precios iguales al costo marginal. Recordemos el ejemplo del puente. La preocupación por un sistema de precios en el que los usuarios dejen de pagar los costos fijos del puente es que tal sistema genera señales falsas acerca de la asignación eficiente de los recursos. Dado que los viajeros usan el puente antes que la barcaza o transbordador, se supone que los recursos debieran trasladarse de la construcción de barcazas o transbordadores a la construcción del puente; pero los viajeros no prefieren el puente porque sus costos sociales sean menores sino porque no pagan sus costos sociales plenos. Sin embargo, los ferrocarriles eran una industria declinante en el momento de la decisión de los moldes de lingotes y un desplazamiento de algunas transacciones de las líneas de lanchones a los ferrocarriles no habría inducido

[21] American Commercial Lines *vs.* Louisville & N.R.R., 392 U.S., 571 (1968). Desde entonces, las enmiendas a la Ley de Comercio Interestatal, que culminaron con la aprobación en 1980 de la Ley Staggers Rail de la fijación de precios no uniformes, les han dado a los ferrocarriles una flexibilidad mucho mayor en la fijación de sus precios.

una expansión ferroviaria económicamente injustificada: sólo habría asegurado la utilización máxima de los activos ferroviarios.

La sospecha de la dependencia reguladora acerca de los alejamientos de la regla del precio igual al costo medio se ha debido, en gran medida, a que son detonados a menudo por la amenaza o la realidad de la competencia. Si una empresa regulada sirve a dos mercados y no afronta la competencia en ninguno de ellos, y el volumen y la elasticidad de la demanda de sus servicios son iguales en ambos mercados y el costo marginal del servicio a los clientes es también el mismo en ambos mercados, se maximizará la producción (sin un déficit) cobrando el mismo precio en cada mercado. Este precio será igual a los costos medios totales (es decir, fijos más variables) de la empresa. Si en uno de los mercados surge un competidor, esto volverá más elástica la demanda afrontada por la empresa regulada. La respuesta eficiente de ésta, aunque parece discriminatoria, consiste en bajar su precio en el mercado competitivo y aumentarlo en el mercado monopólico. Supongamos que los costos fijos de la empresa ascienden a 100 dólares, su costo marginal (= costo variable medio, pues supongamos por el momento que su costo marginal es constante) es de 2 dólares, y el número de unidades demandadas en cada mercado es 50 (para un total de 100), haciendo los costos totales en ambos mercados iguales a 300 dólares (100 fijos más 200 variables). Antes de que surja el nuevo miembro, el precio será de 3 dólares. Supongamos que el neófito tenga costos fijos de sólo 40 dólares y un costo marginal de 2 dólares, de modo que se imagina que puede arrebatarle el mercado a la empresa existente cobrando un precio ligeramente menor a 3 dólares, lo que todavía le rendirá una ganancia sustancial. La empresa existente puede conservar este mercado bajando su precio allí a 2 dólares al mismo tiempo que aumenta su precio en su mercado monopólico a 4 dólares hasta que el miembro recién ingresado se aleje, como lo hará finalmente. Si el neófito sólo puede cubrir sus costos marginales a corto plazo, irá a la quiebra cuando su planta existente se agote. Los costos fijos son, en realidad, parte de los costos marginales a largo plazo del negocio, suponiendo que la continuación del negocio requerirá el remplazo de la planta existente a costos similares a los incurridos en su construcción.

El efecto de la respuesta de la empresa existente al nuevo miembro será que cargue todos los costos fijos a los clientes del mercado en el que la empresa tenga un monopolio. Esto podría parecer "injusto" para ellos y para el neófito. Pero, si la empresa regulada se viera obligada a mantener un precio de 3 dólares en el mercado competitivo, perdería ese mercado. Para permanecer en la industria tendría que aumentar de todos modos a 4 dólares el precio en su mercado monopólico, y los costos sociales totales del servicio a ambos mercados aumentarían de 300 a 340 dólares (es decir, por el monto de los costos fijos del recién ingresado).

Al suponer que la empresa regulada *tiene* costos fijos, no necesitamos suponer que tiene un monopolio natural o siquiera, como el ejemplo numérico

haría esperar naturalmente, que los costos fijos son una fracción sustancial de sus costos totales. Cuando una empresa vende en más de un mercado y tiene costos que son comunes a ambos mercados, tales como los gastos generales corporativos, entonces, *por lo que se refiere a cada mercado*, éstos son costos fijos porque no varían con los cambios de la producción vendida en ese mercado. El ejemplo convencional del economista, de la fijación del precio para cubrir los costos comunes, no tiene nada que ver con el monopolio natural; es el ejemplo humilde de la venta de cueros y carne del mismo animal.[22] Los costos principales de ambos productos son comunes a ambos, y la producción conjunta de los dos productos se maximiza con precios que asignen los costos comunes inversamente a la elasticidad de la demanda, porque entonces se minimiza el efecto del costo (cuando se traslada a un precio) en la reducción de la demanda.

XIII.6. La televisión de paga

La prolongada controversia sobre si debiera permitirse la televisión de paga, una controversia resuelta ahora a favor de este servicio, constituyó una variante del problema del puente y la barcaza o transbordador.[23] El argumento económico en contra de la televisión de paga es que, dado que el costo marginal de la transmisión a otro televidente es cero una vez que el transmisor está en operación, los televidentes debieran estar en posibilidad de comprar programas de televisión a un precio de cero, como ocurre en el sistema de la televisión gratuita (para el televidente). Pero, cuando el costo marginal es cero, el problema del financiamiento de las producciones cuyo precio es igual al costo marginal se agudiza, como hemos visto. La solución tradicional de la industria televisiva consistía en la venta de tiempo de televisión a los anunciantes antes que a los televidentes. Esto permite que los costos de la difusión sean sufragados sin un subsidio gubernamental y sin alejarse del principio del precio igual al costo marginal en el mercado del televidente. Pero, dado que la televisión como un medio de publicidad y promoción de ventas tiene sustitutos bastante buenos, y dado que la demanda del producto anunciado por parte de un televidente no se correlaciona con el placer que obtiene del programa, los anunciantes no pagarán más de unos cuantos centavos por televidente

[22] Los economistas distinguen entre los "costos conjuntos" y los "costos comunes", empleando el primer término para referirse a los casos, tales como el del cuero y la carne, donde los dos productos se elaboran en proporciones fijas. Para nuestros propósitos puede omitirse este refinamiento terminológico.

[23] Véase un animado debate sobre esta cuestión en Jora R. Minasian, "Television Pricing and the Theory of Public Goods", 7 *J. Law & Econ.,* 71 (1964); Paul A. Samuelson, "Public Goods and Subscription TV: Correction of the Record", 7 *J. Law & Econ*., 81; James M. Buchanan, "Public Goods in Theory and Practice: A Note on the Minasian-Samuelson Debate", 10 *J. Law & Econ*., 193 (1967); Samuelson, "Pitfalls in the Analysis of Public Goods", 10 *J. Law & Econ*., 199; Minasian, "Public Goods in Theory and Practice Revisited", 10 *J. Law & Econ*., 205.

por tiempo de televisión. Esto vuelve imposible que se paguen los costos de una programación cara, a menos que pueda reunirse un auditorio de decenas de millones de personas. Los anunciantes no apoyarán un programa de ópera cuya producción cueste 400 000 dólares si atrae a un auditorio nacional de sólo un millón, porque entonces se estarán pagando 40 centavos de dólar por televidente, y podría llegar a los consumidores a un costo menor por otros medios. Pero un millón de televidentes podrían estar dispuestos a pagar 40 centavos de dólar por cabeza para presenciar la ópera. Así pues, la ausencia de la televisión de paga los obligaría a trasladarse a un entretenimiento menos deseado o más costoso.

XIII.7. Tributación por regulación (subsidio interno o subsidio cruzado)

En el caso de los moldes de lingotes, la estructura ineficiente de las tarifas subsidiaba a un grupo de competidores. En otros casos, tales estructuras de tarifas subsidian a clases particulares de clientes. A veces las empresas reguladas proveen efectivamente el servicio a precios por debajo del costo marginal; así ocurría con el servicio ferroviario interurbano de pasajeros antes de Amtrak. Pero aun los precios por arriba del costo marginal podrían ser demasiado bajos desde el punto de vista de la eficiencia. Supongamos que una empresa tiene costos fijos de 500 dólares y un costo marginal (constante) de 1 dólar, y que si fija su precio igual a su costo total medio podrá vender 1 000 unidades a 1.50 dólares por unidad, mientras que si utiliza un sistema de precio de dos partes, en cuyos términos contribuye cada cliente con 10 dólares a los costos fijos pero luego podría comprar cuantas unidades quiera por 1 dólar cada unidad, podrá vender 2 000 unidades. El sistema de dos partes favorece al comprador grande; por ejemplo, el precio medio para un comprador de 100 unidades es de 1.10 dólares, mientras que bajo el sistema de precio igual a costo medio pagaría 1.50. Pero quien compre sólo una unidad se encuentra en peor situación. Este comprador paga 11 dólares en lugar de 1.50. Estaría mejor bajo un sistema del precio igual al costo medio, aunque ese sistema fuera ineficiente porque induciría a los consumidores dispuestos a pagar entre 1 y 1.49 dólares a consumir más bien otros productos cuya producción le cuesta a la sociedad más de 1 dólar. Antes de la desregulación de la industria ferroviaria, los productos agrícolas retribuían menos que la porción de los costos fijos que en justicia (es decir, de acuerdo con la intensidad de su demanda) les corresponde, debido a la persistencia del sistema de fijación del precio por el valor del servicio en una época de vigorosa competencia camionera. El resultado fue un aumento de la asignación del costo fijo a bienes, tales como las manufacturas, para los que había modos de transportación que eran buenos sustitutos, con el resultado de que los embarcadores de tales bienes abandonaron la transportación ferroviaria.

En ambos ejemplos —el precio por debajo del costo marginal y el hecho de que no se logren concentrar los costos fijos en los clientes a los que se les puede hacer pagarlos—, un grupo de clientes resulta en efecto gravado para que pague un subsidio en beneficio de otro grupo. La prohibición regulatoria tradicional de la "discriminación personal" (cobrar un precio diferente a clientes distintos por el mismo servicio… lo mismo desde el lado de la demanda, no de la oferta, del mercado) es otro ejemplo de tributación por regulación.[24] Dado que es más barato atender a los usuarios de instalaciones de alta densidad que a los usuarios de instalaciones de baja densidad, se debería cobrar menos a los primeros por el mismo servicio, pero la norma contra la discriminación personal prohíbe esto.

La analogía de la tributación pone de relieve la naturaleza esencialmente pública de la transferencia de ingresos generada por el sistema de precios. Una empresa no regulada no vendería por debajo del costo marginal, excepto en el caso raro de la rebaja depredadora del precio. Los ferrocarriles habrían abandonado el servicio de pasajeros mucho tiempo antes del surgimiento de Amtrak si los abandonos no hubieran requerido el permiso de la Comisión de Comercio Interestatal. Ni una empresa no regulada utilizaría el precio al costo medio cuando pudiera aumentar sus ganancias con alguna versión del precio de Ramsey.

Es fácil condenar por ineficiente al subsidio interno de lo que yo llamo "tributación por regulación", porque genera precisamente la clase de sustituciones ineficientes que hemos identificado en la discusión de las consecuencias de los controles incorrectos de ganancias. Pero, si suponemos por el momento que es loable el objetivo del subsidio interno y luego nos preguntamos cuáles otros métodos están disponibles para proporcionar el subsidio, veremos que la condena es superficial. Por ejemplo, si el dinero para el subsidio se obtuviera aumentando la tasa del impuesto federal al ingreso, esto sería ineficiente en el mismo sentido en que el subsidio interno es ineficiente (compárese antes III.6). Porque así como el subsidio interno hace que el valor de un producto parezca mayor de lo que es al deprimir artificialmente sus precios, la tributación del ingreso hace que el ocio, y el ingreso no pecuniario tal como las ganancias imputadas de un ama de casa, parezcan más valiosos de lo que son al gravar solamente a sus sustitutos (véanse más adelante XVIII.6 y XVIII.7).

Sin embargo, el subsidio interno puede ser un método de tributación peculiarmente ineficiente. Al exigir que las empresas reguladas mantengan elevadas diferencias de precio y costo en algunos mercados a fin de sufragar el costo del servicio de subsidio en otros mercados, alienta la entrada ineficiente. Supongamos que el costo medio de la empresa regulada en un mercado es de 2 dólares, pero su precio es de 3 dólares, no porque sus ganancias no estén efectivamente controladas por la dependencia reguladora sino porque la em-

[24] Richard C. Levin, "Regulation, Barriers to Exit, and the Investment Behavior of Railroads", en *Studies in Public Regulation*, 181 (Gary Fromm [comp.], 1981).

presa se ve obligada a vender el mismo servicio de 2 dólares en otro mercado (del mismo tamaño) a un precio de sólo 1 dólar. Una empresa que pudiera servir al primer mercado a un costo de 2.50 dólares tendría un incentivo para entrar a ese mercado, aunque el resultado fuera un costo del servicio mayor. A fin de impedir el desperdicio de recursos que implicaría tal entrada (correctamente llamada "llevarse la flor y nata, lo mejor"), así como el derrumbe del programa de subsidios, la dependencia reguladora deberá controlar la entrada. Esto sería completamente innecesario si no existiera el subsidio interno, pero podría eliminarse, sin acabar con el subsidio interno, estableciendo un impuesto indirecto (explícito) sobre las ventas del primer mercado y destinando las recaudaciones a las empresas que sirvan al segundo mercado. La entrada ya no tendría que estar regulada, excepto que cada empresa que entrara al mercado de precio elevado estaría sujeta al impuesto indirecto. La ventaja ineficiente del nuevo miembro en nuestro último ejemplo quedaría eliminada. Y el subsidio interno requiere el control regulador de la salida al igual que de la entrada; de otro modo, la empresa regulada abandonaría simplemente los servicios que la dependencia quiere proveer a tarifas no remunerativas.

Sin embargo, el subsidio interno es a menudo difícil de distinguir empíricamente de la fijación eficiente de los precios, sobre todo en los servicios de red tales como la telefonía (véase antes X.11). La adición de un nuevo suscriptor confiere un beneficio a los suscriptores existentes, así que a fin de obtener el número de suscriptores correcto debiera cobrarse más a los suscriptores existentes y menos a los suscriptores nuevos, quizá menos que el costo marginal. Pero, si la compañía telefónica hace esto, como cuando no cobra una prima a los suscriptores ubicados lejos del intercambio local más cercano, parecerá que está subsidiando a tales suscriptores con los ingresos obtenidos de otros suscriptores. El mercado de costo bajo resultará atractivo para quienes buscan llevarse lo mejor, que estarán gorroneando a costa del transportador telefónico común. La resistencia a esta colada de lo mejor se asemejará a la defensa de un subsidio interno, pero podría ser en efecto la defensa de un sistema de precios eficiente, es decir, uno que permita interiorizar los beneficios externos.

XIII.8. La demanda de regulación

Las deficiencias de la regulación de los servicios públicos considerada como un método para la regulación de las ganancias, el grado en el que parece conservar deliberadamente estructuras tarifarias ineficientes, y la frecuencia con la que se ha impuesto en industrias naturalmente competitivas y se ha empleado también para desalentar la competencia en industrias que tienen cierta cantidad (pero no un número predominante) de características de monopolio natural (como los ferrocarriles, por ejemplo), podrían llevarnos a

preguntarnos si el verdadero propósito de la regulación de las empresas de servicios públicos es el de responder a la preocupación del economista por las consecuencias ineficientes de los monopolios naturales no regulados. Quizá la regulación se vea más bien como un producto, muy semejante a otros, excepto los provistos por el gobierno, demandado por grupos políticos de presión y suministrados a ellos, antes que al público consumidor en general. Los consumidores particulares podrían demandar una estructura tarifaria que, aunque sea ineficiente en general, les otorga beneficios mayores que los costos que les impone en común con otros consumidores. Los miembros de una industria competitiva se benefician de la imposición de controles a los servicios públicos porque la regulación de la tarifa mínima provee una seguridad mayor de un precio efectivo de cártel que el acuerdo privado —al mismo tiempo que se ubica al cártel fuera del alcance de las leyes antimonopólicas—, y el control regulado de la entrada puede eliminar una de las principales amenazas para el éxito de un cártel. Las coaliciones formadas por grupos de consumidores de intereses especiales (como los embarcadores de ciertos productos) y los miembros de una industria podrían ser especialmente eficaces para manipular el proceso regulador.[25]

En la segunda parte explicamos el derecho común como un sistema diseñado (en general) para promover la eficiencia. Aquí empleamos el análisis económico para refutar la concepción de que otra rama del derecho, la regulación de los servicios públicos, se muestra igual de consistente al perseguir la eficiencia, y para proponer, en efecto, que a menudo persigue un objetivo conflictivo. En el capítulo XX haremos un esfuerzo para explicar la diferencia (ilustrada además por los materiales del capítulo siguiente).

XIII.9. El movimiento de desregulación; la privatización

El término "movimiento de desregulación" se refiere a la remoción o reducción de los controles *comprensivos* de industrias particulares, las clases de controles, en las clases de industrias (principalmente de transportación, incluidas la aviación, los camiones de carga, los ferrocarriles, los lanchones, los ductos y el servicio de taxis; de comunicaciones, incluidas la telefonía y sus contrapartidas de datos y la televisión, en particular la televisión por cable; de energéticos, especialmente la energía eléctrica y el gas natural; y las instituciones financieras, primordialmente los bancos pero también sustitutos de los bancos tales como las asociaciones de ahorros y préstamos) que se discuten en este capítulo. El movimiento, que se puede decir que inició a fines del decenio de 1970

[25] Véanse, por ejemplo, Richard A. Posner, "Taxation by Regulation", 2 *Bell J. Econ. & Mgmt. Sci.,* 22 (1971); Thomas W. Gilligan, William J. Marshall y Barry R. Weingast, "The Economic Incidence of the Interstate Commerce Act of 1887: A Theoretical and Empirical Analysis of the Short-Haul Pricing Constraint", 21 *RAND J. Econ.,* 189 (1990).

en la industria aeronáutica, aunque hubo anticipaciones anteriores, coincidió con el aumento de la regulación en los mercados de la salud, la seguridad y el trabajo, de modo que resulta equívoco hablar de "desregulación" en sentido amplio; la regulación ha cambiado en lugar de disminuir.

La sección anterior de este capítulo examinó la demanda de regulación. ¿Pero qué produjo la demanda de desregulación que se tradujo en un movimiento exitoso? Se pueden identificar tres factores. El primero y quizá el más importante es la innovación que redujo el costo de la competencia en relación con el monopolio. Por ejemplo, el desarrollo del fondo del mercado monetario permitió la competencia con los bancos para los depósitos (véase XVI.13), y los bancos presionaron por la relajación de la prohibición de pagar intereses sobre los depósitos bancarios, una prohibición anticompetitiva típica de la regulación integral, de modo que pudieran hacer frente a la competencia de los fondos del mercado monetario. Del mismo modo, las innovaciones en la telefonía de larga distancia forzaron la apertura del monopolio de AT&T a la competencia, y las innovaciones continuas en las telecomunicaciones han abierto a la competencia toda la gama de servicio. Un patrón estándar de desregulación es el aumento de la competencia no regulada, lo que lleva a las empresas reguladas a exigir que se relaje el control regulatorio para que puedan cumplir con la nueva competencia.

El segundo factor era la inflación crónica de la década de 1970. Las agencias regulatorias se tardaron en responder a las solicitudes de las empresas reguladas de que se les permitiera elevar las tarifas, y muchas de estas empresas decidieron a la larga que se encontrarían mejor sin la regulación tarifaria, aunque de esa forma se abriría paso a la fijación competitiva de los precios. Y en tercer lugar, hay una tendencia natural de que la regulación tipo los cárteles resultaría contraproducente. Los límites al ingreso de las aerolíneas en pares de ciudades, por ejemplo, las llevaron a competir por el tráfico que esos límites volvieron lucrativo por la competencia de un servicio caro (como incrementar el número de vuelos entre cada par de ciudades y proporcionar desayunos elaborados e incluso entretenimiento durante el vuelo) que consumieron las ganancias que habían generado las restricciones a la entrada antes de que las ganancias mermaran por la competencia de servicios.

La desregulación ha sido parcial con frecuencia, especialmente en la regulación financiera y en las telecomunicaciones. El sistema de AT&T se desintegró a principios de la década de 1980, pero las compañías operadoras de Bell, que habían sido los principales componentes del sistema, mantuvieron sus monopolios locales como empresas independientes. De hecho, hasta la llegada del teléfono celular, el servicio telefónico local siguió siendo un monopolio natural. La Ley de Telecomunicaciones de 1996 (que modifica la Ley de Comunicaciones de 1934) obliga a las compañías telefónicas locales a proporcionar a las empresas de telecomunicaciones rivales un "acceso desagregado" a algunas de sus instalaciones, entre otras disposiciones. Supongamos

que la compañía local (L) tiene una instalación de conmutación para enrutar las llamadas telefónicas a su destino, y una compañía de telecomunicaciones rival (T) quisiera enrutar las llamadas de sus propios clientes a través de esa instalación. Supongamos que la instalación tiene un gran exceso de capacidad y, por lo tanto, podría encargarse del tráfico de T a un costo mínimo y que para T sería prohibitivamente costoso construir su propia instalación porque no tendría suficiente tráfico para poder recuperar su inversión. Entonces, si L se negara a conceder a T acceso a su servicio de conmutación a un costo cercano al costo de L, T no podría competir.

Pero supongamos en cambio que el mercado para el servicio de T es lo suficientemente grande como para permitir a la compañía recuperar el costo de invertir en su propia instalación de conmutación. T seguiría prefiriendo ir a cuestas de la instalación de L, con la esperanza de que los reguladores hicieran que L cobre un precio tan bajo que estaría subvencionando a su competidor. La Ley de Telecomunicaciones conduce hacia un curso intermedio, al exigir acceso desagregado a costo sólo a los servicios de red que el operador solicitante necesita para poder atender a sus clientes. Pues mientras los operadores solicitantes recurran a los servicios de red suministrados por las compañías telefónicas locales, la competencia se obstaculiza porque los servicios siguen siendo monopolios y requieren regulación.[26] Limitar el requisito de acceso desagregado a costo a los servicios de red que los operadores solicitantes necesitan en vez de sólo querer sirve para "alejar [a esos operadores] de la dependencia respecto de elementos de red desagregados para que se construyan redes terrestres totalmente competitivas, ahora que existe un acuerdo generalizado de que el servicio local ya no es un monopolio natural".[27] Así, la ley de 1996 creó un marco para la desregulación gradual de la industria a medida que los avances en la tecnología y la expansión de los mercados crean un margen mayor de competencia con las compañías telefónicas locales; el acceso desagregado al costo de la empresa local no es necesario en los grandes mercados porque éstos tienen suficiente tráfico para permitir que las compañías de telecomunicaciones rivales financien sus propios centros de conmutación.[28]

Se debe diferenciar entre los efectos económicos y los financieros de la regulación (o bien, entre los efectos de eficiencia y de distribución). Los primeros son los efectos sobre la asignación de los recursos. Los últimos son los efectos sobre la riqueza relativa de las empresas reguladas, sus proveedores, competidores, clientes, etc., y se ilustran por el fenómeno de los "costos encallados". Supongamos que una empresa eléctrica de servicio público ha invertido cantidades enormes de capital en instalaciones de generación de energía nuclear o de otra clase con la esperanza de poder recuperar su inversión en las

[26] Véase Graeme Guthrie, "Regulating Infrastructure: The Impact on Risk and Investment", 44 *J. Econ. Lit.*, 925 (2006).

[27] Illinois Bell Telephone Co. *vs*. Box, 526 F.3d, 1069, 1073 (7lh Cir., 2008).

[28] *Ibid.,* 533-536.

tarifas permitidas por la comisión del servicio público. Supongamos ahora que se permite la entrada competitiva. El nuevo miembro, a menos que tenga costos hundidos comparables a los de la empresa ya existente, no tendrá ninguna inhibición con la fijación de precios por debajo de las tarifas existentes a fin de capturar una porción del mercado. Al actuar así, podría impedir que la empresa ya establecida recuperara su inversión y aun podría mandarla a la quiebra. ¿Debiera impedirse esto prohibiendo la entrada hasta que se haya amortizado la inversión? El economista se inclina a decir "no", que los apuros financieros de los accionistas de la empresa regulada no constituyen un argumento en contra de la utilización más eficiente de los escasos recursos económicos de la sociedad que la competencia hace posible. Pero en contra de esto puede argumentarse que el rembolso de los costos encallados podría ser el precio para superar la oposición de la industria regulada a la desregulación.

Los efectos más importantes de la desregulación se manifiestan en la estructura de la industria (número y tamaño de las empresas), sus costos (incluidos los costos de la mano de obra), los patrones de fijación de precios, la innovación, la variedad de productos o servicios y las actitudes gerenciales. Al eliminarse la regulación amplia se expone a la empresa regulada a la competencia y la nueva entrada, creando presión en favor de lo que en el contexto europeo se llama "racionalización", es decir, a favor de la organización de la industria según el patrón más eficiente. Esto podría significar más empresas, menos empresas, una integración más vertical, menos vertical; lo importante es sólo que una industria cuya estructura es óptima a la luz de la regulación amplia a la que está sujeta probablemente no será óptima en un ambiente competitivo. Se ha estimado que la racionalización ha disminuido, en un periodo de tres años, los costos de la industria ferroviaria en 3 000 millones de dólares, y los de la industria camionera en más de 20%.[29] Una razón por la que la desregulación provoca una baja en los costos es que los sindicatos tienden a ser poderosos en industrias en las que la entrada de empresas nuevas se limita (de hecho, suele evitarse por completo) mediante la regulación. Cuando se desregula una industria sindicalizada, los salarios se reducen porque esa industria es particularmente vulnerable a la entrada de empresas no sindicalizadas, ya que tienen una ventaja en cuanto a costos.[30]

La apertura de una industria que antes estaba comprensiblemente regulada a la entrada no puede dejar de afectar a la estructura de los precios (apar-

[29] Véase Curtis Grimm y Robert J. Windle, "Regulation and Deregulation in Surface Freight, Airlines and Telecommunications", en *Regulatory Reform and Labor Markets,* 15, 24, 28 (James Peoples [comp.], 1998). Por lo que toca a los efectos benéficos de la desregulación de las instituciones financieras, véase Jith Jayaratne y Philip E. Strahan, "Entry Restrictions, Industry Evolution, and Dynamic Efficiency: Evidence from Commercial Banking", 41 *J. Law & Econ.,* 239 (1998).

[30] Véase James Peoples, "Deregulation and the Labor Market", *J. Econ. Perspectives*, verano de 1998, p. 111.

te de los cambios de precios resultantes de la reducción de los costos totales, como los costos salariales), a menudo dramáticamente, como ha ocurrido en las industrias de aeronáutica y de telecomunicaciones. Recuérdese que sin control sobre la entrada resulta insostenible un patrón en el que algunos precios son menores al costo y otros suficientemente mayores al costo para financiar las pérdidas resultantes; no se da un equilibrio. Los nuevos miembros buscarán los mercados de precios altos, obligando a la empresa ya existente a bajar sus precios en tales mercados y, a fin de evitar las pérdidas generales en consecuencia, a elevar sus precios en los mercados en los que había estado cobrando precios menores al costo. En las telecomunicaciones, la apertura ha conducido a precios menores para las llamadas telefónicas de larga distancia y las llamadas comerciales, y a precios mayores para el servicio residencial local; la industria había mantenido muy bajos esos precios a fin de generar un apoyo político generalizado a la estructura reguladora, incluyendo las barreras de entrada, que favorecía la industria. En la industria aeronáutica la desregulación ha significado precios menores para las distancias largas y precios mayores para las distancias cortas. En virtud de que la desregulación implica precios mayores al igual que menores, resulta difícil determinar el efecto neto sobre los niveles de precios; pero las pruebas, consistentes con la teoría, indican que tales niveles han bajado en términos reales (es decir, ajustados por la inflación y ajustados por la calidad) debido a la desregulación.

El efecto sobre la estructura de los precios deja en claro que la desregulación afecta a la distribución de la riqueza y a la riqueza agregada. Aun si los consumidores en conjunto están en mejor situación a resultas de la desregulación, algunos grupos particulares de consumidores —quienes disfrutaban tarifas menores al costo subsidiadas por tarifas monopólicas para otros consumidores— podrían experimentar una disminución de su bienestar neto.

Otro efecto económico de la desregulación puede haber sido la alteración de las mentalidades gerenciales[31] y, por lo tanto, de la variedad y la innovación de productos y servicios. Una empresa ampliamente regulada tiende a ser políticamente sensible y estar guiada por un plan en el sentido de subordinar la respuesta a los deseos cambiantes de los consumidores al interés en el desarrollo sostenido, pronosticable de la empresa, ya que la dependencia reguladora podría tardar en permitir que la empresa cambiara sus precios y productos en respuesta a las fluctuaciones rápidas de las demandas de los consumidores. Una empresa competitiva se preocupa poco por la política y, al ser impulsada

[31] Grimm y Windle, nota 29 anterior, en 25, descubren, por ejemplo, que con la desregulación del transporte ferroviario bajó la edad media de los administradores, así como sus años de servicio con el ferrocarril particular que los empleaba, y que aumentaron sus años de educación formal, todo lo cual implica un desplazamiento hacia un grupo gerencial más flexible, menos escurridizo y más educado. Los jóvenes se sienten más cómodos con el cambio que los viejos, y el capital humano general (creado, en gran medida, por la educación) equipa a una persona para afrontar el cambio mejor que el capital humano específico de la empresa.

por el imperativo competitivo de dar al consumidor lo que desea cuando lo desea, está comprometida a responder con agilidad ante los cambios en la demanda. Cuando AT&T estaba en su apogeo, aunque era tecnológicamente progresista, también era famosa por su insensibilidad a las variadas necesidades de empresas y consumidores. Se concentraba en llevar a cabo reducciones de costos sostenidas en un paquete de servicios muy estandarizado, lo que le permitía cosechar ganancias manteniéndose un paso adelante de los reguladores, quienes fijan los precios sobre la base de los costos pasados antes que de los costos actuales y futuros. La desregulación reveló una heterogeneidad enorme de las demandas de servicios de telecomunicaciones, a las que una industria competitiva nueva respondió con imaginación y entusiasmo. La dirección de la innovación cambió dramáticamente: de la reducción del costo de los servicios existentes a la creación de servicios nuevos.

En un contraste superficialmente paradójico, la industria aeronáutica ha presenciado la "mercantilización asequible" del viaje aéreo. En la era regulada de la competencia de precios restringida, las aerolíneas luchaban entre sí para ofrecer un servicio lujoso, glamoroso —simbolizado por los bares con piano que American Airlines instaló en sus aviones Boeing 747—, que atrajera a los clientes de otras aerolíneas. La competencia de precios que resultó de la desregulación identificó rápidamente una enorme demanda desatendida de servicios sin oropeles, para cuya satisfacción se movilizaron las aerolíneas. El resultado ha sido un aumento enorme del volumen del tráfico aéreo, aunado a una disminución sustancial de los precios reales. Se ha estimado que para 1993 habían bajado en promedio las tarifas de las aerolíneas 20% (en términos reales, es decir, ajustados por la inflación), a causa de la desregulación.[32]

La desregulación no genera cambios automáticos en el comportamiento de la empresa. Cambia los incentivos afrontados por los administradores, y éstos difieren en su capacidad para responder inteligentemente a dichos cambios. Las empresas más viejas, las empresas que eran rentables antes de la desregulación y las empresas familiares tienden a mostrarse más lentas en su respuesta a los desafíos de la desregulación que las empresas que tienen los atributos contrarios.[33] Particularmente interesante, en vista de las problemáticas relaciones raciales de los Estados Unidos, ha sido el hallazgo de que la desregulación disminuye la discriminación racial en las industrias anteriormente reguladas, presumiblemente al aumentar el retorno a las prácticas de contratación por méritos.[34]

[32] Steven A. Morrison y Clifford Winston, "The Fare Skies", *Brookings Review*, otoño de 1997, p. 42. Hay una evaluación fuertemente positiva de las ganancias económicas generales derivadas de la desregulación en Clifford Winston, "Economic Deregulation: Days of Reckoning for Microeconomists", 31 *J. Econ. Lit.,* 1263 (1993).

[33] Morrison y Winston, nota 32 anterior, en 166-167.

[34] Véase John S. Heywood, "Regulated Industries and Measures of Earnings Discrimination", en *Regulatory Reform and Labor Markets*, nota 29 anterior, en 287.

Con todo, la desregulación no se puede declarar un éxito total; la desregulación de los servicios financieros fue demasiado lejos y preparó el camino para el colapso financiero de 2008 (véanse XVI.10 y XVI.14 más adelante).

La privatización es a las empresas de propiedad pública —que en los Estados Unidos incluyen escuelas y universidades, cárceles y prisiones, aeropuertos y control aéreo, ferrocarriles de pasajeros y suburbanos, servicio de pronóstico del tiempo, recolección municipal de basura, hospitales, sistemas de seguros de pensiones e incapacidades (notablemente el Seguro Social) y compañías de energía eléctrica— lo que la desregulación es a las empresas reguladas. Lo que se trata de saber es si el gobierno tiene una ventaja comparativa en cualquiera de estas actividades. En algunas de ellas, como la generación de energía eléctrica, la respuesta es claramente negativa.[35] Es más, podemos formular una declaración mucho más categórica: las empresas de propiedad estatal tienden a ser mucho menos eficientes que las empresas privadas con fines de lucro porque las primeras están controladas por burócratas que carecen de los incentivos y restricciones normales para aumentar al máximo las ganancias, además de verse movidos por grupos de interés, como los sindicatos, que carecen igualmente de esos incentivos y restricciones. (No es que las limitaciones que ejercen los mercados de productos y de capital sobre la gestión empresarial garanticen un comportamiento eficiente. Véase el capítulo XV después.)

Por lo que toca a algunos servicios, hay un argumento poderoso en el sentido de que el mercado privado no proveería la cantidad óptima del servicio (como la educación), pero no está claro por qué el gobierno tiene un papel operativo y no simplemente de subsidio al sector privado. Sin embargo, en el caso de las cárceles y prisiones, puede sostenerse que, en virtud de que el costo del servicio es mucho más fácil de determinar y monitorear que la calidad del servicio, la privatización se traduciría en una declinación de la calidad porque el contratista privado contrarrestará las dramáticas reducciones de costos con reducciones ocultas de la calidad.[36] Lo que importa saber entonces es la importancia relativa asignada a la calidad y al costo. Entre más difícil es para el principal vigilar al agente u operador, los costos operativos o de mediación son mayores; y es más difícil vigilar la calidad de la producción del agente que la cantidad.

Para completar el círculo, si el gobierno posee un monopolio natural, ¿existe aún un argumento a favor de la privatización, dados los problemas que hemos visto de regular los monopolios naturales? Sin embargo, quizás existan sustitutos contractuales eficaces para regular el monopolio. En el próximo capítulo se estudia esta posibilidad, terminando en una discusión extensa de la economía de la privatización.

[35] Véase un resumen de las pruebas en Andrei Shleifer y Robert W. Vishny, "A Survey of Corporate Governance", 52 *Journal of Finance,* 737, 767-768 (1997).

[36] Véase Oliver Hart, Andrei Shleifer y Robert W. Vishny, "The Proper Scope of Government: Theory and an Application to Prisons", 112 *Q. J. Econ.,* 1127 (1997).

Lecturas recomendadas

Armstrong, Mark, y David E. M. Sappington, "Regulation, Competition, and Liberalization", 44 *J. Econ. Lit.*, 325 (2006).

Baumol, William J., y David F. Bradford, "Optimal Departures from Marginal Cost Pricing", 60 *Am. Econ. Rev.,* 265 (1970).

Coase, Ronald H., "The Marginal Cost Controversy", 13 *Economica*, 169 (1946).

———, "The Economics of Uniform Pricing Systems", 15 *Manchester School of Econ. & Soc. Stud.*, 139 (1947).

———, "The Federal Communications Commission", 2 *J. Law & Econ.,* 1 (1995).

———, "The Theory of Public Utility Pricing and Its Application", 1 *Bell J. Econ. & Mgmt. Sci.*, 113 (1970).

———, John C. Panzar y Robert D. Willig, *Contestable Markets and the Theory of Industry Structure*, caps. 1, 12, 16 (ed. rev. 1988).

———, y J. Gregory Sidak, *Toward Competition in Local Telephony* (1994).

Kaserman, David L., y John W. Mayo, "Cross-Subsidies in Telecommunications: Roadblocks on the Road to More Intelligent Telephone Pricing", 11 *Yale J. Regulation,* 119 (1994).

Posner, Richard A., "Taxation by Regulation", 2 *Bell J. Econ. & Mgmt. Sci.,* 22 (1971).

Regulation and Economics (Roger J. Van den Bergh y Alessio M. Pacces [comps.], 2012).

Sappington, David E. M., y Dennis L. Weisman, *Designing Incentive Regulation for the Telecommunications Industry* (1996).

Spulber, Daniel F., *Regulation and Markets* (1989).

Teske, Paul, Samuel Best y Michael Mintrom, *Deregulating Freight Transportation: Delivering the Goods* (1995).

Train, Kenneth E., *Optimal Regulation: The Economic Theory of Natural Monopoly* (1991).

Problemas

1. Entre otras reglas que gobiernan a los servicios públicos y que hemos discutido en este capítulo, existen las siguientes: un servicio público no podrá negarse arbitrariamente a servir a un cliente, no podrá cobrar un precio diferente del que aparezca en una lista de precios publicada (tarifa) y no podrá emprender la construcción de una planta adicional sin el permiso de la dependencia reguladora. ¿Puede explicar la función de estas reglas? ¿Según la teoría de que el propósito de la regulación de los servicios públicos es la promoción de la eficiencia económica? ¿Según la teoría de que su propósito real

es subsidiar a grupos de clientes particulares? ¿Según la teoría de que su propósito real es el de facilitar los precios monopólicos por medio de empresas reguladas?

2. La mayoría de los hospitales estadunidenses ahora son empresas sin fines de lucro. Se cree generalmente que estos hospitales están mal administrados. Incurren en déficit enormes y, según se dice, están aquejados de una capacidad excedente crónica. Tratando de resolver estos problemas, se ha propuesto que los hospitales se conviertan en empresas de servicio público sujetas a los controles habituales de los servicios públicos, en particular al control de las construcciones nuevas (véase el problema 1 anterior). ¿Qué le parece esta propuesta? ¿Sería la nacionalización de la industria hospitalaria mejor o peor desde un punto de vista económico? Véase *Regulating Health Facilities Construction* (Clark C. Havighurst [comp.], 1974); Clark C. Havighurst, "Health Planning for Deregulation: Implementing the 1979 Amendments", *Law & Contemp. Probs.,* 33 (invierno de 1981).

3. ¿Puede haber un problema de alentar la entrada ineficiente a un mercado de monopolio natural si el monopolista emplea el precio en múltiples partes? ¿Si discrimina en el precio?

4. Una ley estatal prohíbe que la autoridad aeroportuaria estatal cobre tarifas irrazonables. En el único aeropuerto internacional del estado, la autoridad (propietaria del aeropuerto) cobra cuotas de aterrizaje a cada una de las aerolíneas con base en la cantidad relativa de espacio del aeropuerto que utilice la aerolínea, y las cuotas se calculan de tal modo que su total sea igual a los costos totales de la operación del aeropuerto, incluida la depreciación, excepto por los costos de las varias concesiones del aeropuerto. Además, la autoridad cobra rentas a los estacionamientos, las empresas de renta de automóviles y otros concesionarios, y estas cuotas generan ingresos sustancialmente mayores que los costos de las concesiones para el aeropuerto. ¿Está la autoridad violando la ley estatal?

5. La Ley Staggers Rail faculta a la Comisión de Comercio Interestatal para exentar de la regulación de la tarifa máxima a cualquier bien acerca del cual no haya peligro de que los ferrocarriles abusen de su poder de mercado (si lo tienen). Supongamos que el precio del carbón estadunidense vendido para exportación se determina por un mercado mundial del carbón, y que el precio mundial de mercado no se ve afectado por la cantidad de carbón estadunidense que se vende en él. ¿Se sigue de aquí que, si los ferrocarriles convinieran en no competir entre ellos en cuanto a las tarifas que cobren por la transportación del carbón a los puertos, para su venta en el mercado de exportación, no tendrían poder monopólico sobre la transportación del carbón, y sus tarifas estarían efectivamente constreñidas por el precio del mercado mundial del carbón? Véase Coal Exporters Assn. of U.S., Inc. *vs*. United States, 745 F.2d, 76 (D.C. Cir., 1984). (Supóngase que no hay competencia intermodal, es decir, proveniente de medios de transporte distintos de los ferrocarriles.)

6. También en los términos de la Ley Staggers Rail, la Comisión no puede invalidar una tarifa para un servicio ferroviario por considerarla irrazonablemente elevada mientras los ingresos generados por la tarifa no excedan 180% de los costos variables del servicio. Véase 49 U.S.C., 10709. ¿Cuál es la justificación económica de esta regla, y es esta regla sensata?

7. ¿En cuáles circunstancias, si las hay, debiera permitirse que un monopolista natural incluya los gastos de publicidad en su costo del servicio?

8. Respondiendo al rápido crecimiento de los gastos federales con el programa Medicare, en la Ley de Reducción del Déficit de 1984 el Congreso impuso un congelamiento por 15 meses sobre el rembolso gubernamental de los honorarios cobrados por los servicios de los médicos bajo Medicare y también sobre el monto de tales honorarios. Explique los efectos probables de estos congelamientos sobre la demanda y la oferta de servicios de médicos, los ingresos de los médicos, el bienestar de los receptores de Medicare y el déficit federal.

9. La Comisión Federal de Energéticos (FPC), predecesora de la Comisión Federal Reguladora de la Energía (FERC), solía regular los precios cobrados por los productores de gas natural. A fin de estimular la producción continua a pesar de los controles de precios, la Comisión permitía que los productores cobraran precios mayores por el gas proveniente de pozos de reciente perforación. No estaba dispuesta a permitirles que cobraran el mismo precio por el gas proveniente de pozos antiguos, por temor a que los productores obtuvieran ganancias súbitas (los pozos antiguos ya habían sido pagados y habían sido perforados a un costo menor). Supongamos que el precio del gas proveniente de pozos nuevos es de 1 dólar por cierta unidad de producción estándar, el precio del gas proveniente de pozos antiguos es 50 centavos de dólar y se producen cantidades iguales en los pozos nuevos y los antiguos, de modo que el precio medio del gas es de 75 centavos de dólar. La mayoría de las ventas son de mezclas del gas proveniente de pozos nuevos y antiguos, de modo que 75 centavos de dólar es lo que pagan la mayoría de los compradores. Este método de fijación del precio elimina las ganancias súbitas. ¿Es un método eficiente?

XIV. LA ELECCIÓN ENTRE LA REGULACIÓN Y EL DERECHO COMÚN

XIV.1. La regulación óptima

El monopolio, la contaminación, el fraude, el error, la mala administración y otros desafortunados efectos colaterales del mercado se consideran convencionalmente como fallas de los mecanismos que lo autorregulan y, por ende, como motivos que harían conveniente someterlo a la regulación pública. Esta percepción del asunto es equívoca. La falla es de ordinario una falla del mercado *y* de las reglas del mercado prescritas por el derecho común. La contaminación, por ejemplo, no se consideraría como un problema grave si los remedios del derecho común, tales como el perjuicio y el allanamiento, fuesen todos eficientes para la minimización de los costos de la contaminación. La elección se plantea raras veces entre un mercado libre y una regulación pública. Se plantea entre dos métodos del control público —el sistema del derecho común de los derechos que se hacen cumplir privadamente y el sistema administrativo del control público directo— y debiera depender de una ponderación de sus ventajas y desventajas en contextos particulares.

El método del derecho común recurre mínimamente a los funcionarios públicos —los jueces y otros empleados judiciales— y básicamente a los ciudadanos privados: a saber, las víctimas y sus abogados; y crea incentivos para la obediencia al amenazar a los victimarios con tener que compensar a las víctimas por el daño que se les causa al violar las reglas. La regulación directa recurre mucho más a los funcionarios públicos (el personal de la agencia reguladora) y trata de impedir que ocurran perjuicios por principio de cuentas, en lugar de compensar a las víctimas de los perjuicios. Con el derecho común del perjuicio, el contaminador podría ser obligado, en un juicio, a pagar una compensación a las personas perjudicadas por la contaminación; con las leyes federales que rigen sobre el aire limpio, el agua limpia y la contaminación, las dependencias públicas crean y administran los estándares destinados a impedir la emisión de contaminantes en cantidades nocivas.

Estas distinciones simples nos permiten pronosticar cuándo es probable que el método del derecho común sea inferior a la regulación directa. Si el perjuicio infligido por una actividad a cada una de las víctimas es demasiado pequeño para que resulte costeable un juicio, habrá un argumento a favor de la regulación directa, siempre que el perjuicio total infligido por la actividad sea sustancial en relación con el costo de la prevención. (Este argumento no es irrefutable; como veremos en el capítulo XXII, la demanda colectiva es un

instrumento para la agregación de reclamaciones pequeñas hasta integrar un solo juicio legal grande.) Quizá de manera sorprendente, el argumento a favor de la regulación directa recupera terreno cuando el perjuicio no es muy pequeño sino muy grande. Un victimario podría no tener recursos suficientes para pagar un juicio de daños muy grandes, y en tal caso su incentivo para cumplir con la ley podría reducirse por debajo del nivel óptimo. Podríamos pensar que mientras la condena se llevara toda la riqueza del victimario, proveería un incentivo suficiente. No es así (véase antes VII.2). Supongamos que, en la Fórmula de Hand, *B* es 100 dólares, *P* es 0.001 y *L* es un millón de dólares, de modo que *PL* = 1 000 dólares, pero la condena máxima que podría pagar el demandado (potencial) sería 10 000 dólares. Entonces, al decidir si gastará *B,* el demandado estará comparando un gasto de 100 dólares con un costo esperado de la condena de sólo 10 dólares (10 000 dólares × *P*).

El derecho común tiene un problema en el caso de las lesiones fatales que se relaciona con su problema en el caso de los perjuicios muy grandes en general, aunque no es idéntico. Se relaciona porque una muerte es una clase de lesión particularmente costosa, que podría agotar los recursos del victimario, pero es diferente debido a que resulta difícil asignar a la vida humana un valor en dólares. Aunque pareciera que este problema no podría evitarse sustituyendo el derecho común por la regulación directa de la seguridad —que los reguladores, al determinar cuán estricta debe ser la regulación de la seguridad, tuvieran que asignar por lo menos un valor implícito a la vida humana—, esto no es cierto. La interrogante económica planteada por una propuesta, digamos, de obligar a los ferrocarriles a remplazar con luces titilantes las señales de cruces en los cruceros no es el valor que el viajero asigna a su propia vida en un cruce ferroviario, sino el valor que asigna a una mayor reducción de un riesgo que ya es pequeño. Tenemos bastante información acerca de cómo evalúa la gente las medidas de seguridad que reducen el riesgo de lesión o de muerte; el problema para el sistema de los cuasidelitos consiste en extrapolar desde estas estimaciones para llegar al valor de la vida misma. Aunque en el capítulo VI sugerimos un método para hacer esto, es limitado, como vimos, y el sistema de los cuasidelitos no lo ha considerado. Sin embargo, el argumento a favor de la regulación se debilita si se basa en un defecto del sistema de los cuasidelitos que puede corregirse.

Otro problema de la regulación del derecho común, examinado también en el capítulo VI, surge del hecho de que la relación causal existente entre un victimario particular (o aun una clase de victimarios) y una víctima particular podría atenuarse. Si tenemos una idea muy precisa de que un accidente en un reactor nuclear causará un aumento de 0.01% en el número de los cánceres, pero no sabemos cuáles cánceres no habrían ocurrido si no se hubiese producido el accidente, será difícil imponer los costos del accidente al propietario del reactor nuclear a través de los métodos del sistema de cuasidelitos, a no ser, quizá, sino mediante demandas colectivas que amalgamen las

reclamaciones del daño probabilístico de todas las personas que habían estado expuestas al peligro en cuestión.

La contaminación del aire plantea este problema y otro: el de la incertidumbre acerca de la identidad del victimario. Una lesión o perjuicio dados (como una enfermedad pulmonar, ropa sucia o hedores), derivada de la contaminación del aire, se deberá a menudo a las emisiones combinadas de numerosos contaminadores, y podría resultar imposible agregar todos los contaminadores en un solo juicio legal, o establecer una conexión causal entre un contaminador dado y la lesión. Sin embargo, estos problemas aconsejan la regulación directa complementaria, en lugar de preventiva. No hay problema en combinar el derecho común con la regulación directa (como con la conducción) siempre y cuando las obligaciones impuestas no sean inconsistentes.

De modo que el derecho común es un sistema regulador imperfecto; pero también lo es la regulación directa. Tiende a ser más costosa que la regulación del derecho común, porque es continua; la maquinaria del derecho común se invoca sólo si alguien resulta efectivamente lesionado. (Adviértase la analogía existente con la interrogante de si deberá remunerarse a los rescatadores —restitución— o castigarse a quienes no rescaten —responsabilidad cuasidelictiva—. Véase antes VI.9.) La regulación directa tiende también a politizarse más que la regulación del derecho común porque los jueces están mejor protegidos que los administradores contra la recompensa y la retribución políticas. Y aunque los organismos reguladores están especializados y disponen de medios más flexibles para obtener información que los tribunales, también enfrentan serios problemas de información: si las víctimas de accidentes no tienen nada que ganar al llamar la atención del gobierno sobre una condición insegura, los reguladores pueden tener dificultades para descubrir cuál es el problema.

La regulación funciona mejor cuando la manipulación de algunos insumos de la seguridad o la salud bien conocidos pueden generar resultados justificados por los costos. Un buen ejemplo es el de castigar a quienes conducen en estado de ebriedad. Los costos externos superan casi seguramente a los beneficios del conductor, y la dificultad de medir los costos de los accidentes fatales aconseja que se trate de impedir la ocurrencia de los accidentes prohibiendo la conducta peligrosa antes de que se produzca uno.

Como en estos ejemplos, la regulación tiende a usar medios de control *ex ante* (preventivos), mientras que los litigios del derecho común tienden a usar medios *ex post* (disuasivos). El primer método lo ilustran los límites de velocidad, el segundo las demandas de lesiones por negligencia. Pero, como en este ejemplo, los dos tipos de regulación suelen unirse. La regulación de la seguridad en las carreteras es un mosaico complejo de regulaciones *ex ante* (incluidos los límites de velocidad y otras reglas de seguridad, estándares federales diseñados para la seguridad, estándares para el diseño y mantenimiento de carreteras, así como la licencia de los conductores) y regulaciones

ex post (como las demandas por conducción imprudente, demandas por responsabilidad del producto debido a defectos en el diseño o fabricación de vehículos de motor y procesos penales por ebriedad u otra conducción imprudente).

Sin embargo, la regulación *ex ante* puede ser tanto judicial como administrativa, como en la prisión preventiva, los requerimientos y los decretos regulatorios, y la regulación *ex post* puede ser administrada por dependencias y tribunales, como la Comisión Federal de Comercio y la Junta Nacional de Relaciones Laborales, que funcionan principalmente mediante procedimientos de prueba realizados después de haber ocurrido una violación de las leyes aplicadas por la dependencia.

El enfoque *ex ante* fomenta la claridad de la obligación jurídica y, por lo tanto, presumiblemente un mejor cumplimiento (menos violaciones inadvertidas) al establecer reglas antes del inicio de las actividades reguladas. A diferencia de una demanda, la regulación *ex ante* se activa antes de que haya una pérdida; puede diseñarse e imponerse de manera central (por ejemplo, por una sola dependencia, como la Administración de Alimentos y Medicamentos, en lugar de un sistema judicial descentralizado); y puede hacerse cumplir mediante sanciones leves, porque la sanción óptima para la creación de un simple riesgo de lesión suele ser menos draconiana que la sanción óptima por provocar una lesión real. Sin embargo, esto significa que la regulación *ex ante* y la *ex post* son en realidad inseparables. Debido a que el cumplimiento de las reglas nunca es de 100%, debe haber una maquinaria para castigar a los infractores, aunque la maquinaria podría implicar sanciones impuestas por la misma dependencia reguladora, con participación judicial limitada a la revisión judicial del procedimiento de sanción. Y, aunque las reglas implican costos fijos altos (el diseño de la regla en primer lugar), si una regla es muy clara y las violaciones de ella se castigan severamente, el cumplimiento puede lograrse sin procedimientos de ejecución frecuentes, por lo que el costo marginal de la regla podría ser bajo. Por lo tanto, las reglas son atractivas cuando la alternativa serían estándares vagos, dando lugar con frecuencia a violaciones que son efectivas o cabe sostener que lo son y, por tanto, a procedimientos de ejecución frecuentes.

Como muestra esta discusión, la regulación *ex ante* y las reglas tienen una afinidad. La regulación *ex ante* permite explotar las propiedades economizadoras de las reglas como preventivas. Con estándares vagos, el énfasis regulador se desplaza hacia la búsqueda de disuasión mediante procedimientos para castigar a los infractores.

La regulación *ex ante* restringe la base de información cuando adopta la forma de reglas porque entonces está comprando precisión al costo de excluir la información específica de cada caso que los promulgadores de la regulación no previeron o descartaron para mantener en forma simple la regulación (es decir, mantenerla como regla). Los estándares, como la negligencia, permiten

considerar más información en casos particulares, aunque al hacerlo no sólo se reduce la previsibilidad, sino que se convierten en regulaciones *ex post* porque los estándares vagos generan disputas que requieren resolverse en litigios. Aun así, la regulación *ex ante*, como la atención preventiva en medicina, puede representar una carga sobre mucha actividad inofensiva, como la conducción segura que excede el límite de velocidad. Comparemos el chequeo de toda la población en busca de afecciones médicas que sólo aquejan a unas cuantas personas.[1]

Cuando la regulación *ex ante* adopta la forma de licencia en lugar de prohibición —compárese el requisito de un permiso de construcción con un límite de velocidad— los costos del cumplimiento podrían aumentar, junto con un mayor riesgo de soborno si la licencia es muy valiosa. En cambio, la regulación *ex post* podría requerir sólo intervenciones escasas (de nuevo, comparemos el chequeo de las condiciones médicas con tratamiento siempre y cuando una condición produce síntomas): cero en el caso limitante en el cual una regla o estándar alcanza el 100% de cumplimiento, aunque existen costos de cumplimiento. La regulación *ex post* economiza en gastos administrativos porque la intervención es esporádica y utiliza tanto información específica de caso como el procedimiento contencioso. Hay más información, incluyendo información actualizada y específica de cada caso, y se examina y se sopesa con más cuidado porque se presenta en un procedimiento impugnado. En su carácter privado (a diferencia del público) y contencioso, el litigio como enfoque regulador toma prestados los métodos de los mercados competitivos.

La ventaja de la información de la regulación *ex post* es particularmente marcada cuando el estándar es la responsabilidad estricta, lo que significa que el causante está obligado a pagar los daños incluso si no hubiese podido evitar ocasionar la lesión a un costo razonable. Un ejemplo es una lesión causada por el uso de explosivos, vista como una actividad extremadamente peligrosa, en la construcción de un túnel. Los victimarios potenciales tienen un fuerte incentivo para equilibrar los costos y los beneficios de la actividad peligrosa con el fin de decidir si, o en qué escala o en qué circunstancias, involucrarse en tal actividad se justifica por el costo, y tienen acceso fácil a la información necesaria.

La regulación *ex post*, caracterizada por la adjudicación del derecho común con su gran énfasis en los estándares, como la negligencia y la buena fe, en lugar de las reglas, implica altos costos por cada caso en comparación con la adjudicación de una multa por exceso de velocidad. Esto se debe, en parte, a la información adicional generada por un procedimiento centrado en una lesión específica provocada en circunstancias particulares. Más información puede no sólo volver más costoso un procedimiento sino también crear más incertidumbre y, como resultado, más diferencias en el resultado; la incertidumbre también hace que sea más difícil vigilar el desempeño del juez u

[1] La elección entre reglas y estándares se revisa más a fondo después en XXI.3.

otro regulador para asegurarse de que sea competente y honesto. Y, un punto relacionado con el hecho de que cuando ha ocurrido una lesión la sanción óptima es mayor que cuando se ha creado un riesgo que aún no se ha materializado, el causante podría no tener recursos suficientes para pagar la sanción.

Entre mayor sea el costo social si fracasa la disuasión y si es más probable que la disuasión fracase, más fuerte será el argumento a favor de la regulación *ex ante*. Incluso si 99% del desplome de la construcción pero sólo 10% de los delitos relacionados con drogas pueden evitarse mediante la regulación *ex post* (demandas por diseño o construcción negligentes en el primer caso, y castigos penales en el segundo), el costo social de 1% del desplome de la construcción podría exceder el costo social de 90% de los delitos relacionados con drogas. Si también supera el costo de la prevención mediante la promulgación y ejecución de códigos de construcción, entonces se justifica la regulación *ex ante*. La correlación positiva entre la gravedad de la lesión y la probabilidad de que la disuasión fracase refuerza esta conclusión. Están correlacionadas positivamente porque es probable que la solvencia limitada de los causantes potenciales haga que *para ellos* el costo esperado de la lesión (recordemos que estamos suponiendo una lesión grave) sea menor que el costo social esperado.

Un ejemplo similar es la inspección pública de los restaurantes y las plantas de procesamiento de alimentos, en lugar de confiar totalmente en la amenaza de demandas por negligencia para evitar la intoxicación alimentaria. Muchos restaurantes son empresas nuevas pequeñas, con dueños y personal inexpertos, sin capital de reputación que se vería afectado por un envenenamiento accidental, y sin recursos financieros que podrían agotarse por una demanda de cuasidelitos. Un régimen de inspección podría de hecho facilitar la entrada de nuevos comensales, ya que en su ausencia muchas personas podrían temer a probar un nuevo restaurante. Y el descuido en una planta de procesamiento de alimentos puede provocar lesiones masivas, cuyo costo para las víctimas podría exceder la capacidad de pago del procesador de alimentos negligente.

XIV.2. Una revisión del fraude al consumidor

Aunque hay ciertas fuerzas del mercado que operan para brindar a los consumidores información acerca de los productos que compran, podrían no funcionar bien siempre, y el recurso del derecho común para remediar el fraude tampoco (véase antes IV.7). Pero no olvidemos el papel de las marcas comerciales en la protección de los consumidores contra los engaños (véase antes IV.7). La inversión que un productor hace en su marca comercial es como un rehén: aumenta el costo, para el productor, de reducir subrepticiamente la calidad del producto o de tratar de engañar de otro modo a los consumidores, porque cuando descubren el engaño el productor podría perder toda su inversión. Eso puede ocurrir aunque no descubran el engaño, porque el fra-

caso al intentar mantener el control de calidad es un obstáculo legal para hacer que un competidor infractor respete la marca comercial.

Dado el papel de las marcas comerciales en la protección de las expectativas de los consumidores, resulta paradójico que hayan sido criticadas porque inducen a los productores a gastar dinero en la creación de una imagen espuria de alta calidad, que permite la obtención de rentas monopólicas induciendo a los consumidores a que se olviden de sustitutos baratos de calidad igual. Consideremos productos como la aspirina y los líquidos blanqueadores hogareños; las marcas comerciales como Bayer y Clorox, aunque son químicamente equivalentes a sus sustitutos genéricos, se venden a precios mayores. Pero el hecho de que dos bienes tengan la misma fórmula química no los hace de igual calidad. Es posible que no se elaboren con la misma atención al control de calidad. El fabricante que ha invertido mucho en una marca comercial tiene un incentivo mayor para el mantenimiento de la calidad y, sabiendo esto, un consumidor racional estaría dispuesto a pagar una prima por la marca de ese fabricante.

Comparemos los remedios públicos administrados por la Comisión Federal de Comercio (FTC) con los remedios del mercado y del derecho común contra el engaño al consumidor. Originalmente, los consumidores tenían escaso incentivo para invocar a la maquinaria de cumplimiento forzoso de la Comisión, la que no podía otorgar el pago de daños a un consumidor defraudado. La amenaza de una queja ante la Comisión induciría a veces a un vendedor a arreglarse con un consumidor enojado, pero, una vez iniciados los procedimientos de la Comisión, el vendedor no tendría ningún incentivo para buscar tal contemporización; esto limitó el interés de los consumidores por presentar quejas ante la FTC. (Sin embargo, los tribunales han sostenido recientemente que la Comisión puede ordenar la restitución a los consumidores defraudados.[2]) Por el contrario, los competidores del vendedor tienen un incentivo constante para quejarse ante la Comisión, quien puede expedir una orden de cesar y desistir para acabar con la práctica que está quitándoles clientes. Pero, en virtud de que la FTC sufraga todos los gastos del procedimiento, el vendedor quejoso no tiene ningún incentivo para abstenerse de plantear quejas esencialmente frívolas destinadas a fastidiar a un competidor antes que a aclarar la mala información para el consumidor.

La combinación de la falta de un incentivo adecuado para que el consumidor se queje con la posesión por parte del competidor de un incentivo demasiado grande crea un desbalance en la naturaleza de las presiones que se hacen sentir sobre la Comisión. Ésta oye menos de los consumidores defraudados que de los sindicatos preocupados por el efecto de las importaciones chinas baratas sobre las ventas de los empleadores de sus miembros, de los peleteros

[2] FTC *vs*. Verity Int'l, Ltd., 443 F.3d, 48,66-67 (2nd Cir., 2005); FTC *vs.* Garvey, 383 F.3d, 891, 900 (9th Cir., 2004); FTC *vs*. Think Achievement Corp., 312 F.3d, 259, 262 (7th Cir., 2002).

preocupados por la competencia de las pieles sintéticas y de los comerciantes de diamantes preocupados por la competencia de los diamantes sintéticos. Dada la naturaleza de estos insumos, no resulta sorprendente que, a través de su historia, gran parte de las reglas y decisiones de la Comisión se hayan relacionado tenuemente, si acaso, con el problema real del engaño al consumidor.[3]

Además, la Comisión carece de las armas necesarias para ser eficaz contra el operador pirata, quizás el mayor defraudador de los consumidores. Individuo poco fiable, se aprovecha de la gente que tiene menores probabilidades de quejarse ante la Comisión, y de todos modos ésta carece de sanciones apropiadas para enfrentar a las empresas que carecen de continuidad en su operación, que ocultan sus actividades o que son financieramente irresponsables. La FTC podría convertirse en una dependencia más eficaz para luchar contra el fraude al consumidor. Podría otorgársele poder para imponer sanciones penales, para evaluar los costos del juicio contra los competidores quejosos, y para conceder la autonomía a sus oficinas de campo. El efecto de tales cambios sería aproximar más el modelo administrativo de la regulación al modelo del derecho común.

Una sanción interesante que la Comisión sí utiliza es exigir a una empresa, a la que se ha demostrado que ha hecho publicidad falsa que, en su publicidad futura, corrija la impresión engañosa creada por los anuncios anteriores. Exigir tal "publicidad correctora" es penal —es decir, impone al infractor un costo mayor al daño causado por la infracción—, porque los pronunciamientos podrían hacer que los consumidores desplazaran todas sus compras hacia los competidores, no sólo la parte de sus ventas que se habría perdido en beneficio de los competidores si no hubiese habido publicidad falsa. Por ejemplo, la FTC exigió a Johnson Products Company, fabricante del "relajante capilar" Ultra Sheen, que incluyera en cada empaque un aviso detallado de que el uso del relajante capilar podría provocar quemaduras en el cabello y el cuero cabelludo, pérdida de cabello y lesión ocular. Parece ser que los relajantes capilares fabricados por otras empresas tenían riesgos similares; pero no se les impusieron órdenes por más de un año y, en consecuencia, Johnson perdió ventas considerables. La pérdida se entiende fácilmente cuando consideramos que el consumidor del relajante capilar debía escoger entre un producto que continuamente lo prevenía de su peligro y los productos que parecían más seguros porque no contenían ninguna prevención. Una sanción penal con estos efectos es social y privadamente costosa; castiga el engaño pasado a costa de crear nuevos engaños.

[3] Richard A. Posner, *Regulation of Advertising by the FTC* (American Enterprise Institute, 1973); Ross D. Petty, "FTC Advertising Regulation: Survivor or Casualty of the Reagan Revolution?", 30 Am. Buss. L. J., 1 (1992); James S. Wrona, "False Advertising and Consumer Standing under Section 43(a) of the Lanham Act: Broad Consumer Protection Legislation or a Narrow Pro-Competitive Measure?", 49 *Rutgers L. Rev.*, 1085, 1146-1154 (1995).

El problema más general de tratar de proteger a los consumidores de la publicidad falsa es que la publicidad se dirige a una audiencia que es inevitablemente heterogénea en lo tocante a los costos de absorción de la información. Un anuncio que engaña a algunos consumidores omitiendo una reserva importante para ellos podría engañar (o dejar de informar) a otros consumidores si se añadiera la reserva. Supongamos que un anuncio afirma que la cámara del anunciante es más barata que las cámaras de sus rivales, lo que es cierto, pero omite añadir que es también menos duradera. Algunos consumidores saben esto sin que se lo digan; otros no lo saben. Si el anuncio se considera engañoso, el anunciante podría optar por no decir nada acerca del precio. En ese caso, los consumidores que saben que la cámara del anunciante es menos duradera que las cámaras de sus rivales, pero no saben que es más barata, se verán perjudicados.

XIV.3. La revelación obligatoria

Leyes reglamentarias como la Ley de la Veracidad en los Préstamos, e iniciativas de la Comisión Federal de Comercio como las que requieren que se publiquen clasificaciones numéricas en octanos y se revele el contenido de alquitrán y nicotina de los cigarros, utilizan un enfoque distinto para el problema de información al consumidor sobre el producto: el de exigir que los vendedores provean a los consumidores de información que se considere valiosa, en lugar de que simplemente se prohíban las presentaciones engañosas.[4] La revelación obligatoria,[5] como se llama a veces a este enfoque, difiere no sólo de la publicidad correctora sino también del requerimiento de la revelación afirmativa cuando, en su ausencia, el consumidor supondría lo contrario de manera incorrecta (por ejemplo, que el aceite reprocesado era nuevo); la revelación podría exigirse aunque los vendedores no hubiesen sido acusados de hacer presentaciones engañosas.

Como vimos en el capítulo IV, los mercados podrían dejar de proveer de información suficiente acerca de una característica del producto que es, a la vez, costoso de averiguar para el consumidor y común a todas las marcas del producto. La característica podría ser buena o mala desde el punto de vista de un consumidor: el bajo contenido de colesterol de la margarina o el alto contenido de colesterol de la mantequilla. Un productor de margarina podría resistirse a publicitar el bajo contenido de colesterol de su producto porque su

[4] Véase Richard A. Posner, "The Federal Trade Commission's Mandated-Disclosure Program: A Critical Analysis", en *Business Disclosure: Government's Need to Know,* 331 (Harvey J. Goldschmid [comp.], 1979). Otros ejemplos importantes son los avisos requeridos por la Administración de Alimentos y Fármacos (FDA) en las etiquetas de los productos alimenticios y en las etiquetas e insertos de los productos farmacéuticos de los pacientes.

[5] Es también un principio fundamental de la regulación de los valores que se discute después en XVI.11.

publicidad beneficiará a sus competidores, quienes no han ayudado a costear su gasto. Y ningún productor de mantequilla tiene incentivo alguno para publicitar el alto contenido de colesterol de su producto.

El problema no se limita a las marcas que son idénticas en lo tocante a la característica en cuestión. Supongamos que, antes de que se conocieran generalmente los peligros del hábito de fumar, una marca de cigarros contuviera menos alquitrán y nicotina que las demás. Antes de publicitar este hecho, el productor tendría que ponderar el aumento de su participación en el mercado derivado de convencer a los consumidores de que su marca era menos dañina contra la pérdida de sus ventas resultante de informar a los consumidores de cigarros acerca de los peligros del hábito de fumar, induciendo así a algunos de ellos al consumo de otros productos completamente diferentes. Si bien los vendedores de productos competidores habrían tenido un incentivo para alarmar a los consumidores y alejarlos del tabaquismo, no se podría esperar realmente que iniciaran una campaña para que dejaran de fumar. ¿Qué pensarían los consumidores si un fabricante de goma de mascar publicitara que ésta, a diferencia de fumar cigarros, no causa cáncer de pulmón?

Los fabricantes de automóviles tardaron en ofrecer cinturones de seguridad, temiendo sin duda que estarían publicitando los peligros de manejar… mientras que, a la vez, cobrarían un precio más alto para cubrir el costo de dichos artículos.

En 2008 la ciudad de Nueva York comenzó a hacer cumplir una ordenanza dirigida a las cadenas de comida rápida para publicar en los menús el número de calorías de cada elemento, y con el mismo tamaño de letra que el usado para el platillo. El propósito declarado de la ordenanza es reducir la obesidad, más que simplemente ayudar a las personas a evitar que se vuelvan obesas de manera inadvertida. La ordenanza puede tacharse de innecesaria, porque la información acerca del contenido calórico puede comunicarse sin necesidad de que se imprima en letras grandes en el menú (una alternativa sería la publicación en el sitio web de la cadena, o la publicación de un aviso separado en el restaurante), y de paternalista, porque las personas preocupadas por su peso tienen el incentivo y la capacidad para informarse por sí mismas sobre el número de calorías que consumen.

Sin embargo, una ley que pretende reducir la obesidad es paternalista sólo si la obesidad no produce costos externos. Pero sí los causa. Las personas obesas consumen una cantidad desproporcionada de recursos médicos, por los que, por lo general, no pagan totalmente debido a que hay un subsidio considerable a los gastos médicos públicos y privados (privados a través de grupos de seguros que no pueden o se rehúsan a identificar y rechazar a los asegurados de alto riesgo). Sin embargo, la magnitud de la externalidad está en duda, porque las personas obesas mueren, en promedio, a una edad más joven que las personas delgadas y, por lo tanto, consumen recursos médicos por menos años que las personas delgadas.

Como en el caso de los cigarrillos, la competencia no producirá la ingesta calórica deseada por los consumidores dada la renuencia de las empresas a anunciar la seguridad *relativa* porque alerta al consumidor sobre el peligro existente.

La importancia de la ordenanza de la ciudad de Nueva York reside en el requisito de que los números de calorías se impriman junto a los alimentos en los menús en letras grandes. Más que informar, el propósito es alarmar. Los psicólogos han demostrado (lo que de todos modos es bastante obvio) que la gente responde más a la información que se le presenta de una forma dramática y memorable que a la información que se le presenta de forma abstracta o que simplemente se recuerda en lugar de lanzarla a la cara.

El efecto directo de la ordenanza sobre la obesidad puede ser pequeño, pero impondrá costos de cumplimiento en las cadenas de comida rápida sujetas a ella y, como resultado, hará que el precio de la comida rápida aumente, aunque tal vez por una cantidad trivial… y el aumento en el precio contribuirá, aunque de manera modesta, a los esfuerzos para reducir la obesidad.

Se ha sugerido prohibir la publicidad televisiva de alimentos que engordan en programas dirigidos a niños y adolescentes, pero en realidad podría ser contraproducente. Usualmente, la publicidad tiene un mayor efecto en las acciones de la marca que en la demanda agregada de un producto, lo que significa que podría ser mejor para los anunciantes en su conjunto que se prohibiera hacer publicidad. (La publicidad que simplemente altera la participación de mercado de los competidores aumenta los costos para ellos en su conjunto sin beneficiarlos.) Con ganancias mayores, y eliminada la costosa forma de competencia sin precios, los anunciantes podrían competir más en el precio, dando como resultado precios más bajos. Una medida más efectiva para reducir la obesidad juvenil podría ser la prohibición de la venta o servicio de refrescos y otros alimentos con alto contenido calórico en las escuelas, o incluso imponer impuestos elevados a tales alimentos.

Algunos observadores atribuyen parte de la crisis económica que inició con el colapso financiero en septiembre de 2008 a la revelación inadecuada del costo de los productos financieros de consumo. De hecho, durante mucho tiempo se ha revelado obligatoriamente dicho costo. En particular, la Ley de la Veracidad en los Préstamos[6] exige la revelación uniforme de las condiciones de crédito al consumidor, en particular de la tasa de interés. Sin embargo, el valor de tal revelación es cuestionable. Los educados y los ricos pueden protegerse por sí mismos.[7] Y en cuanto a los consumidores que están constreñidos por

[6] Véase a este respecto, por ejemplo, William C. Whitford, "The Function of Disclosure Regulation in Consumer Transactions", 1973 *Wis. L. Rev.*, 400, 420.

[7] Véanse algunas pruebas de apoyo en Thomas A. Durkin, "Consumer Awareness of Credit Terms: Review and New Evidence", 48 *J. Bus.*, 253 (1975); George G. C. Parker y Robert P. Shay, "Some Factors Affecting Awareness of Annual Percentage Rates in Consumer Installment Credit Transactions", 29 *J. Fin.*, 217, 223 (1974). Por lo que toca al derecho y la economía de la regulación

la liquidez y que, por lo tanto, se podría pensar que necesitan y acogerían tal información, en la práctica y de modo realista su elección no se plantea entre comprar bajo un plan de abonos, pedir prestado a un banco y pagar en efectivo, o no comprar sino ahorrar. Su elección se plantea entre planes de abonos. Y el comprador puede comparar la desutilidad de pagar 20 dólares mensuales durante 40 meses con la de pagar 15 dólares mensuales durante 60 meses o 25 dólares mensuales durante 30 meses, sin que se le diga cuáles son las tasas de interés porcentuales anuales implicadas por estos planes de pagos. El comprador escogerá la combinación que sea mejor para él a la luz de su liquidez actual y su liquidez futura esperada, sus gastos actuales y sus gastos futuros esperados, los usos alternativos de su dinero, etc. Estas consideraciones *no* se resumen en una tasa de interés, una cifra que interesa primordialmente a quienes tienen oportunidades de inversión.

No todos los préstamos de personas de recursos modestos toman la forma de préstamos a plazos para permitir la compra de un producto específico. Consideremos los préstamos a corto plazo y con tasa de interés alta. Existen pruebas de que muchas personas no pueden entender las tasas de interés y otros cargos que cobran las compañías de préstamos, en parte debido a la complejidad de la mayoría de los documentos de préstamo. La gente termina pagando tasas ridículamente altas cuando puede pedir prestado a una cooperativa bancaria o de crédito a tasas mucho más bajas o, sin tener que pedir prestado en absoluto, podría recurrir a ahorros que generan un interés bajo. Muchos solicitantes de préstamos, al carecer de autocontrol (es decir, en términos económicos, que tienen tasas de descuento altas) —pero desconocen esto y, por lo tanto, son incapaces de tomar contramedidas—, son incapaces de moderar su deseo de adquirir bienes y servicios y terminan sobreendeudados.[8]

El argumento contrario es que las personas que no pueden pedir prestado a un banco porque tienen mal crédito podrían necesitar un préstamo desesperadamente. Si se coloca un tope a las tasas de interés, estos desafortunados podrían ser incapaces de pedir prestado porque el tope puede ser muy bajo para que la tasa de interés compense al prestamista por el riesgo de incumplimiento. En consecuencia, podrían perder su casa, su automóvil u otra propiedad o encontrarse ellos mismos a merced de los usureros.[9] Es cierto que

del financiamiento al consumo en general, véase Richard Hynes y Eric A. Posner, "The Law and Economics of Consumer Finance", 4 *Am. Law & Econ. Rev.*, 168 (2002). Y por lo que toca al tema general de la competencia del consumidor, véase Chr. Hjorth-Andersen, "The Concept of Quality and the Efficiency of Markets for Consumer Products", 11 *J. Consumer Research*, 708 (1984).

[8] Véase, por ejemplo, Oren Bar-Gill y Elizabeth Warren, "Making Credit Safer", 157 *U. Pa. L. Rev.*, 1, 44-45 (2008); Ronald J. Mann y Jim Hawkins, "Just Until Payday", 54 *UCLA L. Rev.*, 855, 881-882 (2007); Lynn Drysdale y Kathleen E. Keest, "The Two-Tiered Consumer Financial Services Marketplace: The Fringe Banking System and Its Challenges to Current Thinking About the Role of Usury Laws in Today's Society", 51 *S. Car. L. Rev.*, 589, 605-610 (2000).

[9] Véase, por ejemplo, Jonathan Zinman, "Restricting Consumer Credit Access: Household Survey Evidence on Effects Around the Oregon Rate Cap", 34 *J. Banking & Finance*, 546 (2010);

una tasa de interés anual de 300% puede caracterizarse verdaderamente como astronómica. Pero una persona que pide prestados 5 000 dólares a esa tasa y los liquida dos semanas después sólo paga 577 dólares en intereses, y el préstamo podría haberle permitido evitar la ejecución hipotecaria de su casa, o algún otro evento terrible que le habría costado más de 577 dólares. Sin embargo, muchas de las personas que solicitan préstamos a tasas de interés astronómicas terminan apuradas dando vueltas para refinanciar sus préstamos mes a mes, y esto va en contra de la teoría de que éstos son préstamos a corto plazo incurridos racionalmente, a pesar de su elevado costo, como una respuesta temporal a los reveses inesperados.[10]

XIV.4. Alternativas a la revelación: el caso de las grasas trans

Dada la eficacia limitada de la revelación como herramienta reguladora, las prohibiciones absolutas de sustancias o actividades peligrosas tienen un atractivo considerable y, por supuesto, son comunes, y se analizan en la siguiente sección de este capítulo. Un ejemplo inusual, que merece consideración aparte, es otra regulación novedosa de la ciudad de Nueva York: la prohibición de las grasas trans en los alimentos que se venden en restaurantes o en los servicios de banquetes y comidas.

Las grasas trans son, en gran medida, grasas sintéticas muy utilizadas en los alimentos fritos y en productos horneados como conservadores, y existe una opinión médica fuerte de que contribuyen de manera considerable a la enfermedad cardiaca, ya que aumentan el colesterol que es malo para el corazón (LDL) y bajan el colesterol que ayuda a proteger las arterias coronarias contra los efectos del colesterol malo (HDL). Cerca de la mitad de los restaurantes de la ciudad de Nueva York utilizan grasas trans en su cocina (o lo hacían antes de que se aprobara la norma); y aproximadamente un tercio de la ingesta calórica de los neoyorquinos proviene de comidas en restaurantes.

La industria de los restaurantes es altamente competitiva, por lo que, si los consumidores están dispuestos a pagar un precio más alto por las comi-

Todd J. Zywicki, "Consumer Welfare and the Regulation of Title Pledge Lending", Mercatus Center Working Paper No. 09-36 (septiembre de 2009), www.mercatus.org/sites/default/files/publication/WP0936_Consumer_Welfare_and_Regulation_of_Title_Pledge_Lending.pdf (consultado el 4 de diciembre de 2009); Donald P. Morgan y Michael R. Strain, "Payday Holiday: How Households Fare after Payday Credit Bans" (Federal Reserve Bank of New York Staff Reports No. 309, febrero de 2008), http://ftp.ny.frb.org/research/staff_reports/sr309.pdf (consultado el 4 de diciembre de 2009); Mann y Hawkins, nota 8 anterior, en 884-894 (2007).

[10] Véase Michael A. Stegman y Robert Faris, "Payday Lending: A Business Model That Encourages Chronic Borrowing", 17 *Econ. Development Q.*, 8, 19-21 (2003); Drysdale y Keest, nota 8 anterior, en 605-610. Véase más sobre el debate entre estas dos escuelas de pensamiento en Midwest Title Loans, Inc. *vs*. Mills, 593 F.3d, 660 (7th Cir., 2010), y la bibliografía a favor y en contra que se cita en dicho dictamen.

das que no contengan grasas trans, la industria los complacerá. Los restaurantes que atienden a comensales conscientes de la salud anunciarán que no usan grasas trans en la preparación de su comida o indicarán en el menú la cantidad de grasas trans en cada producto. Otros restaurantes atenderán a los comensales que prefieren una comida más barata antes que una más sana. Por lo tanto, la prohibición obliga a las personas que quieren comer en restaurantes a pagar precios más altos, incluso si prefiriesen pagar menos y tomar el riesgo de una mayor probabilidad de enfermedad cardíaca. Algunas de estas personas serían quienes rara vez comen en restaurantes y evitan las grasas trans cuando cocinan en casa, por lo que el riesgo para la salud de una comida de restaurante que contiene grasas trans es menor. Otros serían personas que no creen en la opinión médica —y, a menudo, tal opinión es incorrecta— o piensan que las grasas trans mejoran el sabor de los alimentos o que la prohibición es el resultado de la presión política de los productores de sustitutos de las grasas trans, como el aceite de maíz, o de los restaurantes que han abandonado voluntariamente el uso de grasas trans y no quieren ser puestos en una desventaja competitiva. Dado que la mitad de los restaurantes de Nueva York continúan utilizando grasas trans, esto sugiere que la mayoría de los consumidores no apoyarían la prohibición.

Pero este análisis ignora el costo de asimilar la información, a diferencia del costo de divulgarla. Muchas personas nunca han oído hablar de las grasas trans; muchas que sí han oído no saben que son (muy probablemente) dañinas para la salud; y, sobre todo, casi nadie fuera de las comunidades médicas y de nutrición sabe *qué* tan dañinas son las grasas trans y en qué cantidad y, por lo tanto, cuál sería el beneficio de evitarlas en los restaurantes. Sería imposible explicar el peligro de las grasas trans a un cliente del restaurante, ya que dependería de su consumo diario promedio de éstas, lo que no saben ni él ni el restaurante.

En tal situación, la prohibición de las grasas trans podría producir una mejoría neta en el bienestar de los neoyorquinos. Los defensores de la prohibición estimaron que reduciría en 500 el número anual de muertes por ataque cardiaco en la ciudad de Nueva York. La estimación económica de consenso del valor de una vida estadunidense es de 7 millones de dólares, de modo que un ahorro de 500 vidas produciría un beneficio de 3.5 mil millones de dólares. Los restaurantes incurrirían en algún costo al remplazarlas con otros conservadores, pero es probable que no sea grande. La mitad de los restaurantes de la ciudad de Nueva York ya han eliminado gradualmente las grasas trans sin que nadie notara un gran salto en los precios de esos sitios; y el costo de fabricación de los sustitutos de las grasas trans no parece ser mayor —la única ventaja de las grasas trans es que aumentan un poco el tiempo de caducidad de los alimentos—. Esto es importante para los restaurantes, al permitirles ahorrar en costos de almacenamiento y desperdicio, pero es probable que no sea crítico.

Existen pruebas de que la prohibición, que entró en vigor en 2007, ha sido efectiva para reducir la ingesta de grasas trans de los neoyorquinos.[11]

Una comparación de la revelación obligatoria de información de seguridad con la prohibición de productos peligrosos bien podría favorecer a la prohibición (lo que es, de hecho, en todo caso la solución más común). La revelación obligatoria crea sobrecarga de información; existen tantas advertencias (porque muchas agencias federales y estatales diferentes requieren advertencias sobre tantos productos), y son tan verbosas, que muchos consumidores se desentienden del asunto. Y la redacción exacta suele dejarse a los vendedores, quienes son capaces de rellenar las divulgaciones con información irrelevante que contribuye al efecto de desconexión.[12] La psicología del comportamiento del consumidor (la base de la "economía del comportamiento", véase antes I.4) apoya firmemente la prohibición de productos peligrosos en lugar de tratar de educar a los consumidores a través de la revelación de los peligros. Pero los defensores del enfoque económico-conductual para comprender el comportamiento económico se estremecen ante la idea de abogar por las prohibiciones absolutas, temiendo que la prohibición de productos que son simplemente insalubres etiquetará a los economistas conductuales como "paternalistas" y por ello socaven la aceptación de la economía del comportamiento.[13]

XIV.5. Seguridad y salud

Si no es seguro, un fármaco que se vende a millones de personas puede causar un daño enorme, mientras que los casos particulares de negligencia médica sólo perjudican a un paciente. Por otra parte, es factible probar cada fármaco nuevo y, por lo tanto, determinar su seguridad de antemano, pero es imposible exigir a los médicos que busquen la aprobación de una dependencia reguladora para cada procedimiento que realizan. De acuerdo con este análisis, la regulación *ex ante* es el modo dominante de la regulación de los fármacos nuevos, mientras que la regulación *ex post*, en la forma de juicios por negligencia médica, es el modo dominante de la regulación del tratamiento médico.

Las regulaciones reales de la seguridad y la salud suelen ser estorbosas desde un punto de vista económico. Las leyes que exigen a las compañías farmacéuticas que realicen pruebas prolongadas y caras antes de introducir un fármaco nuevo demoran la introducción de fármacos recientes, lo que se

[11] Véase Alice Park, "NYC's Trans Fat Ban Worked: Fast-Food Diners Are Eating Healthier", *Time Health and Family*, 17 de julio, 2012, http://healthland.time.com/2012/07/17/nycs-trans-fat-ban-worked-fast-food-diners-are-eating-healthier/

[12] Véase Omri Ben-Shahar y Carl E. Schneider, "The Failure of Mandated Disclosure", 159 *U. Pa. L. Rev.*, 647 (2011).

[13] Ryan Bubb y Richard H. Pildes, "How Behavioral Economics Trims Its Sails and Why" (N.Y.U. School of Law, 25 de marzo, 2013).

traduce en la muerte de personas que quizá se habrían salvado con una introducción más temprana. A la luz de estos y otros factores, los estudios de la regulación de fármacos nuevos han descubierto que los costos de la regulación superan a los beneficios.[14]

El carácter errático de la regulación de la seguridad (y de la salud) se demuestra por la variación extraordinaria en las estimaciones del valor de la vida que guían a los reguladores al decidir cuánto obligar a gastar en seguridad a las entidades reguladas. Un estudio reciente concluye que las estimaciones reguladoras del valor de la vida oscilan entre 100 000 y 100 000 millones de dólares por vida salvada.[15] La parte inferior de la gama es demasiado baja, lo que se traduce en una regulación demasiado laxa, y la parte superior de la gama es demasiado alta, lo que se traduce en una regulación que es demasiado estricta. Al requerir una inversión demasiado grande en seguridad, en realidad tales regulaciones pueden matar a más personas de las que salvan. El ingreso y la mortalidad están negativamente relacionados (los ricos viven más que los pobres), y los costos de la regulación operan como un impuesto regresivo porque aumentan los costos marginales de la empresa y, por lo tanto, los trabajadores y los consumidores son, principalmente, aunque no de manera exclusiva, quienes los asumen.[16]

Por otro lado, una regulación que se concentre en un solo insumo de la seguridad podría ser ineficaz o aun nociva. Por ejemplo, los requerimientos de cinturones de seguridad en los automóviles reducen el costo de la conducción rápida para el conductor y sus pasajeros, y sería de esperarse (por la Ley de la Demanda) que los conductores respondieran conduciendo más rápido. De ser cierto, habrá más accidentes, y, si bien los costos de cada accidente para el conductor y sus pasajeros podrían ser menores en virtud de la protección brindada por su cinturón de seguridad, los costos totales de los accidentes para los peatones (quienes no están protegidos por cinturones de seguridad) debieran aumentar, y hay pruebas de que lo hacen.[17] Afortunadamente, parece

[14] Sam Peltzman, "The Benefits and Costs of New Drug Development", en *Regulating New Drugs*, 113 (Richard L. Landau [comp.], 1973); Henry G. Grabowski, "Drug Regulation and Innovation" (Am. Enterprise Institute, 1976); *cf.* Grabowski, John M. Vernon y Lacy Glenn Thomas, "Estimating the Effects of Regulation on Innovation: An International Comparative Analysis of the Pharmaceutical Industry", 21 *J. Law & Econ.*, 133 (1978).

[15] John F. Morrall III, "Saving Lives: A Review of the Record", 27 *J. Risk & Uncertainty*, 221, 230-231 (2003) (cuadro 2). Véase también Cass R. Sunstein, *Free Markets and Social Justice,* 304 (1997) (cuadro 12.2).

[16] *Ibid*., 302-305; International Union *vs*. OSHA, 938 F.2d, 1310, 1326-1327 (D.C. Cir., 1991) (fallo concurrente).

[17] Sam Peltzman, "The Effects of Automobile Safety Regulation", 83 *J. Pol. Econ.,* 677 (1975). Véanse también Steven Peterson, George Hoffer y Edward Millner, "Are Drivers of Air-Bag Equipped Cars More Aggressive? A Test of the Offsetting Behavior Hypothesis", 38 *J. Law & Econ.,* 251 (1995); Theodore E. Keeler, "Highway Safety, Economic Behavior, and Driving Environment", 84 *Am. Econ. Rev.,* 684 (1994); Christopher Garbacz, "More Evidence on the Effectiveness of Seat Belt Laws", 24 *Applied Economics,* 313 (1992). De acuerdo con el análisis de

que a fin de cuentas los cinturones de seguridad crean beneficios netos muy grandes; se ha estimado que su uso salva 15 mil vidas al año.[18] Pero esto crea desconcierto: Dado que los beneficios del uso del cinturón de seguridad por parte de los conductores y los pasajeros exceden los costos en gran medida, ¿por qué razón el mercado libre no indujo la instalación y el uso generalizado de los cinturones de seguridad sin necesidad de la intervención gubernamental? La respuesta podría ser que, si no se educa al público sobre los beneficios de seguridad de los cinturones, la empresa que los ofrezca no podría recuperar el costo en un precio mayor por sus automóviles; y el costo de la campaña de educación en sí sería difícil de recuperar porque las otras empresas podrían aprovecharse gratuitamente de la campaña. En otras palabras, la primera empresa que ofreciera los cinturones de seguridad estaría creando una oportunidad de mercado para las otras empresas sin que a éstas les costara.

XIV.6. Regulación de la contaminación

Hay tres tipos básicos de regulación de la contaminación (aparte de las demandas por perjuicio del derecho común). Un enfoque, el del control de los insumos, consiste en que la legislatura o una dependencia administrativa prescriba las medidas tecnológicas específicas que el contaminador debe tomar a fin de evitar fuertes sanciones legales. Por ejemplo, podría exigirse que un municipio instale cierta clase de planta de tratamiento del drenaje, que una siderúrgica construya sus chimeneas 12 metros más altas, que los fabricantes de automóviles instalen un tipo particular de aparatos de control de las emisiones. Este enfoque requiere que el regulador tenga una cantidad de información enorme acerca de los costos y beneficios de métodos alternativos de control de la contaminación. Una observación relacionada es que la especificación del método particular de control de la contaminación desalienta la búsqueda del método más eficiente. En las deliberaciones ante la legislatura o la dependencia que conducen a la formulación del estándar, la industria afectada tiene un incentivo para proponer el método de control de la contaminación que sea más barato, independientemente de su eficacia, y para negar la existencia o eficacia de métodos más costosos (aunque sean más eficientes). Y, una vez que se adopte la medida especificada, la industria no tendrá ningún incentivo

Peltzman, Philip L. Hersch y Jeffrey M. Netter, "The Impact of Early Safety Legislation: The Case of the Safety Appliance Act of 1983", 10 *Int'l Rev. Law & Econ.*, 61 (1990), concluye que, al requerir frenos mejores, la Ley de Dispositivos de Seguridad alentó a los ferrocarriles a conducir trenes a velocidades mayores y con cargas más pesadas, provocando más accidentes. Pero el aumento de los costos de accidentes exagera los costos sociales de la Ley, porque las velocidades más rápidas y las cargas más pesadas redujeron los costos de envío.

[18] Steven D. Levitt y Jack Porter, "Sample Selection in the Estimation of Air Bag and Seat Belt Effectiveness", 83 *Rev. Econ. & Stat.*, 603 (2001).

para desarrollar métodos mejores, a menos que resulten ser también más baratos.

Un segundo enfoque consiste en establecer el nivel de las emisiones de contaminación que se considera tolerable y obligar luego a los contaminadores, so pena de imponerles una prohibición o una multa, a que no excedan ese nivel, pero dejando la elección del método a la discreción de la industria (control de la producción). La empresa contaminante se verá impulsada a minimizar los costos del cumplimiento del estándar de emisión, lo que es bueno, pero el estándar puede ser ineficiente: puede permitir una contaminación excesiva o demasiado escasa desde un punto de vista económico. El análisis costo-beneficio podría utilizarse para fijar el estándar, pero ello requeriría que la dependencia tuviera tanta información como la empresa acerca de los costos del cumplimiento de diversos estándares. Se eliminaría así la principal eficiencia asociada con la especificación del nivel de las emisiones permitidas antes que de los instrumentos precisos de control de la contaminación que la empresa debe emplear.

Es fácil entender por qué la dependencia regulatoria preferiría fijar un estándar de emisiones sin tomar en cuenta los costos y los beneficios, pero en realidad no se pueden descartar tales consideraciones. Porque, cuando se intenta el cumplimiento del estándar, la empresa o la industria argumentarán que el costo del cumplimiento es prohibitivo, o sea, desproporcionado con los beneficios derivados de la disminución de la contaminación. Tal argumento no puede descartarse a menos que la sociedad desee reducir la contaminación muy por debajo de los niveles eficientes. Y está en juego algo más que la eficiencia. Aun si el abatimiento de la contaminación no es tan costoso que una planta ya no pueda operar económicamente, la limitación aumentará los costos de la planta; y es probable que el propietario responda reduciendo la producción, lo que conducirá a una disminución del empleo y a una baja de los ingresos de los trabajadores. Por supuesto, los efectos distributivos serán los mismos independientemente de que la limitación de la contaminación sea eficiente o no. Una limitación eficiente podría hacer que una planta cerrara; los costos sociales plenos de la planta (incluida la contaminación) podrían superar el valor de su producción a cualquier nivel. Pero, cuando la regulación asume la forma de especificar la tecnología de control de la contaminación particular, las preocupaciones de distribución y eficiencia intervienen en la decisión del regulador en una etapa más temprana.

Un tercer enfoque, que no se ha empleado aún en este país pero es un gran favorito de los economistas, consiste en gravar la contaminación. La tasa impositiva para cada unidad contaminadora se haría igual a los costos sociales estimados que se generan por la contaminación resultante; así se distingue el impuesto de una multa que pretenda impedir la contaminación en la forma habitual de las sanciones penales. Una empresa sujeta a un impuesto a la contaminación compararía sus costos tributarios con los costos

de la compra de equipo de control de la contaminación o de la reducción de su producción o del esfuerzo de disminuir de otro modo la contaminación. Si fuese posible un ahorro tributario neto mediante una de estas medidas, la empresa la adoptaría; en caso contrario pagaría el impuesto y seguiría contaminando.

Este enfoque se asemeja al de la responsabilidad cuasidelictiva estricta (el control de los insumos se asemeja a la responsabilidad por la negligencia y los topes rígidos para la emisión se asemejan a una sanción penal), excepto que se hace cumplir públicamente, antes que privadamente, y no hay defensa de negligencia contribuyente. El impuesto corresponde a los pagos de daños a las víctimas de la contaminación. Se exige que los contaminadores paguen tales "daños", independientemente de que haya o no algunos métodos de control de la contaminación que los evitaran a un costo menor. Esto otorga a los contaminadores un incentivo para buscar y adoptar medidas preventivas de la contaminación justificadas por su costo (incluidos los cambios del nivel de las actividades, como el cierre, la disminución de la producción o la reubicación de sus plantas, quizás a países extranjeros menos preocupados por la contaminación), pero no para adoptar ninguna medida preventiva que cueste más que su beneficio por la reducción de los costos sociales de la contaminación. Se evita así la determinación pública de los niveles de la contaminación justificados por los costos, que es requerida al establecer un estándar de emisión económicamente sensato.

Un impuesto a la contaminación es un ejemplo de una tributación regulatoria, considerada como algo distinto a un impuesto de recaudación. (Los economistas llaman a un impuesto regulatorio un impuesto "pigouviano", en homenaje al economista inglés A. C. Pigou, quien fuera el primero en desarrollar el concepto.) La mayoría de los impuestos, por supuesto, están concebidos para generar ingresos al gobierno; un efecto incidental de ellos es el de reducir la actividad gravada, de manera que, como veremos en el capítulo XVIII, el mejor impuesto recaudador de ingresos es aquel fijado a una actividad elevadamente inelástica, y un impuesto por cabeza es el óptimo en este respecto. El propósito de un impuesto regulatorio puro no es el de recaudar ingresos sino el de desalentar la actividad gravada: en el caso que nos ocupa, la emisión de contaminantes. No obstante, la mayor parte de los impuestos regulatorios generan algún ingreso, ya sea porque están planeados de esa forma o porque el gravamen es demasiado bajo para hacer que evitarlo se justifique por el costo, o bien, un aspecto implícito en el segundo punto, porque las compañías sujetas al impuesto difieren en cuanto a los costos en que tendrían que incurrir para evitarlo.

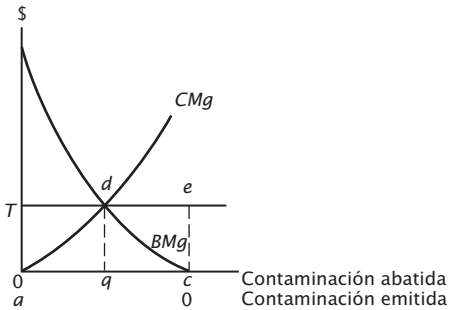
El enfoque fiscal para el control de la contaminación es atractivo, pero no es una panacea:

1. Es probable que resulte contraproducente si la víctima es quien puede evitar la contaminación al costo menor (instalando aire acondicionado, vi-

viendo más lejos de la fábrica, etc.). El contaminador gastará en el control de la contaminación una cantidad igual, en el margen, al ahorro tributario estimado, aunque las víctimas pudieran haber reducido los costos de la contaminación en la misma cantidad a un costo menor.[19]

2. Los accidentes son de ordinario eventos discretos, relativamente infrecuentes. Sus costos pueden evaluarse en procedimientos separados. Pero hay miles de millones de emisiones de contaminantes cada año, y sería imposible estimar el costo social de cada una para fijar las tasas impositivas correctas. Los costos sociales de diferentes contaminantes, y del mismo contaminante en diferentes partes del país (o, para el caso, de un solo estado, condado o aun ciudad o pueblo) no son uniformes. Además, el costo social de la contaminación no es necesaria o probablemente una función lineal de la cantidad de contaminación. Supongamos que cada una de 10 empresas emite 100 unidades de contaminación, y que el costo social agregado es 1 000 dólares. Parecería que el impuesto debiera ser de 1 dólar por unidad, y supongamos que ésa es la tasa que se fija. Entonces, cada empresa buscará medidas para reducir su responsabilidad fiscal, y supongamos que a un costo de 10 dólares cada empresa puede reducir la contaminación de su planta 15%, pero que sería prohibitivo el costo de una reducción mayor. Las empresas incurrirán en este costo y, en consecuencia, habrá 15% menos de contaminación… ¿pero conferiría eso un beneficio de 150 dólares (para un neto de 50 dólares) después de tomar en cuenta su costo? No necesariamente. Es posible que todo el costo social provenga de las primeras 50 unidades de contaminación (5%), y que la contaminación restante no imponga costos incrementales, como podría ocurrir si, por ejemplo, el único daño causado por la contaminación afectara a una especie particular de peces y las primeras 50 unidades fuesen suficientes para matarlos a todos. Entonces el impuesto habrá llevado a las empresas a incurrir en costos de abatimiento que no tienen ningún valor social. O quizá 90% del costo social de la contaminación provenga del último 5% de las emisiones, de modo que si disminuyeran las emisiones 5% sería sólo de 100 dólares el costo social de la contaminación restante. En el primer caso ciertamente, y en el segundo caso posiblemente, el impuesto llevará a las empresas a gastar demasiado en el control de la contaminación. El impuesto correcto sería igual al costo social marginal, antes que medio, de la forma particular de la contaminación, de modo que variaría con el nivel de la contaminación. Pero se

[19] En principio, la víctima de la contaminación podría (y debería) ser también objeto de tributación; la analogía por establecer es con la negligencia contribuyente en un caso de responsabilidad estricta. El impuesto no tiene que ser recaudado; para completar la analogía con la responsabilidad estricta, la responsabilidad fiscal de la víctima sólo tiene que ser deducida de la del contaminante para producir la cantidad económicamente correcta de control de la contaminación, dado que la inhabilidad de las víctimas para obtener la plena compensación (el resultado de la deducción) tendrá el efecto de incentivos deseado. William J. Baumol, "On Taxation and the Control of Externalities", 62 *Am. Econ. Rev.,* 307 (1972). El problema radica en la medida.

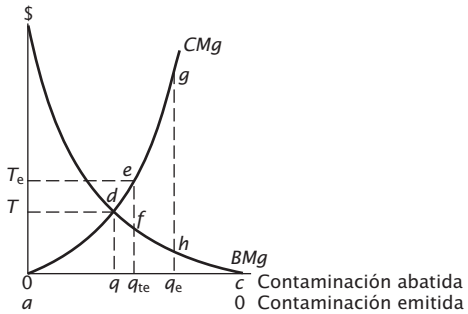


GRÁFICA XIV.1

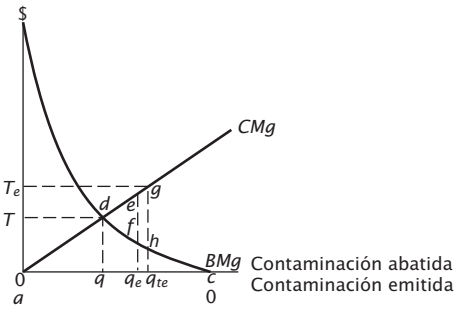
requeriría una cantidad exorbitante de información para elaborar semejante tarifa impositiva.

3. O bien se hará que las empresas sujetas al impuesto lo paguen o bien se inducirá a todas (o a algunas de ellas) a incurrir en costos de control a fin de evitar el gravamen. De cualquier forma, el costo de producción de la empresa se incrementará, aproximadamente en proporción a su producción, y una gran parte del costo más elevado se trasladará a los consumidores de los productos de la firma bajo la forma de precios más altos. Por consiguiente, un impuesto a la contaminación, a semejanza de otros impuestos específicos (véase XVIII.3 después), es regresivo. Así, a fin de asegurar la proporcionalidad o progresividad generales del sistema impositivo, la imposición de impuestos amplios a la contaminación requeriría exenciones, devoluciones o cambios compensatorios en otras partes del sistema impositivo.

En la gráfica XIV.1 se muestran los efectos de riqueza de un impuesto a la contaminación. Se fija el impuesto a un nivel calculado para inducir a la empresa a emitir la cantidad de contaminación óptima: la cantidad en la que los beneficios sociales marginales (*BMg*) derivados del abatimiento de la contaminación son iguales a los costos marginales del abatimiento. Se supone que *BMg* baja a cero (esto es, se interseca con el eje horizontal) en el punto en el que se emite una contaminación de cero. Entonces el punto 0, *a* en el eje horizontal, es la cantidad de contaminación abatida (es decir, 0) cuando no se realiza ningún esfuerzo de abatimiento y la cantidad de contaminación emitida en consecuencia (*a*), mientras que *c*, 0 es la cantidad de abatimiento (*c*) al nivel de contaminación de cero (de modo que $a = c$). En *q*, donde se abate la cantidad de contaminación óptima, la empresa incurre en un costo igual al área *dqa*, bajo *CMg* entre *a* y *q*, y además paga un impuesto, igual al rectángulo *decq*, sobre la contaminación no abatida. Bajo un estándar de emisión que obligara a la empresa a llevar el abatimiento hasta *q*, la empresa incurriría en un costo *dqa* pero no pagaría ningún impuesto. Compárese con la gráfica VI.2 en el capítulo VI; una comparación paralela de la negligencia y la responsabilidad estricta.



GRÁFICA XIV.2



GRÁFICA XIV.3

4. Parecería que un sistema impositivo aplicado a la contaminación le ahorraría al contaminador los costos potencialmente astronómicos del cumplimiento de estándares de emisión excesivamente exigentes; el impuesto sería el tope de la exposición causada por él. Pero, en realidad, resulta imposible decir *a priori* cuál sistema conduce a costos sociales más altos si la regulación (ya sea un impuesto o un estándar de emisión) es demasiado estricta. Esto se muestra en las gráficas XIV.2 y XIV.3. T es de nuevo el impuesto óptimo, y q es el abatimiento óptimo de la contaminación, ya sea inducido por T o prescrito en un estándar de emisión (óptimo). T_e y q_e son, respectivamente, un impuesto y un estándar de emisión más exigentes en la misma proporción. En la gráfica XIV.2, el impuesto induce a la empresa a llevar el abatimiento hasta el punto q_{te}, donde los costos marginales del abatimiento superan a los beneficios marginales, lo que conduce a un desperdicio de recursos medido por el triángulo *def*. El desperdicio es menor que el inducido por el estándar de emisión errónea: el triángulo mayor *dgh*.

Pero en la gráfica XIV.3 se revierte la relación; el impuesto genera el desperdicio mayor. La intuición que se encuentra detrás de estos resultados es que, si los costos marginales del abatimiento aumentan mucho en la región del impuesto, la empresa preferirá pagar el impuesto antes que abatir la contaminación; habrá un pago de transferencia pero no un costo social. Sin embargo, si los costos marginales del abatimiento aumentan más suavemente, la empresa podría preferir llevar el abatimiento muy lejos, lo que se traduciría en un costo social sustancial.

El enfoque del impuesto tiene una gran ventaja sobre los otros enfoques reguladores: no requiere que la dependencia mida los costos del cumplimiento de los criterios de control de la contaminación incorporados en las tasas impositivas (siempre que se omitan los efectos distributivos); la dependencia sólo necesita estimar los beneficios de la reducción de la contaminación. Esto hace que fijar el impuesto al nivel erróneo sea menos probable que cometer ese yerro con un estándar de emisión, caso en el que será me-

nos probable que surja el problema representado en la gráfica XIV.3. Dicho de otro modo, los estándares de emisión requieren un análisis costo-beneficio; los impuestos a la contaminación sólo requieren un análisis del beneficio. Esto significa también que los impuestos a la contaminación generan incentivos para efectuar cambios en el nivel de la actividad y en el nivel del cuidado para reducir la contaminación, puesto que, dando por sentado que la contaminación es más o menos proporcional a la producción, reducir la producción es una manera como puede reducir la empresa su factura tributaria.

Incluso los economistas que simpatizan con las objeciones contra el intento de monetizar los beneficios del abatimiento de la contaminación han sido muy críticos de la Ley del Aire Limpio, el programa más ambicioso del país para el control de la contaminación.[20] Entre los aspectos económicamente más dudosos del programa, algunos de los cuales tienen explicaciones políticas bastante transparentes, se encuentran los siguientes: *1)* la regulación más estricta de las fuentes de contaminación del aire nuevas que de las ya existentes,[21] lo que induce a las empresas a demorar la introducción de tecnologías de producción más limpias, pero también beneficia a las empresas ya existentes al elevar los costos de la nueva capacidad de producción; *2)* la política, favorecida por los productores de los estados de "aire sucio", de no permitir ni siquiera la degradación ligera de la calidad del aire de las áreas más limpias del país, aunque los costos sociales marginales de la contaminación (particularmente para la salud) sean mayores en las áreas sucias que en las limpias, y *3)* la insistencia en que todas las fuentes de contaminación disminuyan sus emisiones, sin tomar en cuenta los diferentes costos marginales del abatimiento entre las fuentes.

En respuesta a las críticas de la tercera característica de la ley, el Congreso autorizó, y la Agencia de Protección Ambiental de los Estados Unidos (EPA) creó un sistema de permisos de contaminación negociables para las emisiones de dióxido de azufre (la causa de la lluvia ácida) generadas por

[20] Véase, por ejemplo, Robert W. Crandall, *Controlling Industrial Pollution: The Economics and Politics of Clean Air,* 57 (1983); Crandall, Theodore E. Keeler y Lester B. Lave, "The Cost of Automobile Safety and Emissions Regulation to the Consumer: Some Preliminary Results", 72 *Am. Econ. Rev. Papers & Proceedings,* 324 (mayo de 1982).

[21] En parte, esto se debe al hecho de que en el caso del control de la contaminación los estándares son más estrictos para las nuevas fuentes de origen que para las ya existentes y, en parte también, al hecho de que el proceso de autorización concedida a las nuevas fuentes se administra con más efectividad que el monitoreo aplicado a las fuentes existentes. La razón de que sean más estrictos los estándares para las nuevas fuentes que para las ya conocidas es de naturaleza política. La imposición en todo punto rigurosa de controles sobre la contaminación sobre una planta ya establecida trae aparejados efectos de desempleo que están focalizados y resultan visibles, mientras que una planta que apenas va a entrar en operaciones todavía no cuenta con empleados, así que no se registran directamente esos efectos. Puede ser que las oportunidades de los desocupados de conseguir un empleo salgan afectadas, pero lo más probable es que este fenómeno quede desapercibido.

las empresas eléctricas de servicio público.[22] Cada permiso (llamado "concesión") autoriza a una empresa de servicio público a emitir una tonelada de dióxido de azufre por año. El número total de las concesiones se ha limitado a un nivel muy por debajo de las emisiones anuales totales de dióxido de azufre de las empresas eléctricas de servicio público del país, de modo que el programa reduce las emisiones totales del contaminante. Pero las empresas quedan en libertad para venderse sus concesiones entre sí, de modo que una empresa que pudiera reducir sus emisiones a costo bajo tendría incentivos para vender algunas de sus concesiones a una empresa que incurriera en un costo elevado para reducir sus propias emisiones, lo que permitiría la reducción de los costos agregados del abatimiento del dióxido de azufre al mismo tiempo que aumenta la cantidad del abatimiento. Tal sistema de control de la contaminación se llama "derechos de emisión" y ha sido sugerido para el dióxido de carbono, la principal causa del calentamiento global. Véase después XIV.10.

XIV.7. El congestionamiento del tránsito

La congestión del tráfico es una externalidad negativa clásica. Un conductor racional egoísta no considera el efecto de su conducción en los otros usuarios de la carretera, sino sólo en sí mismo. El método común para reducir el congestionamiento —construir más carreteras— no sólo es muy costoso, sino que hasta cierto punto es autodestructivo, ya que al reducir el congestionamiento (y, por lo tanto, el costo en tiempo de conducción) se atrae más tráfico. No es de extrañar que, pese a tanta construcción de carreteras, el congestionamiento, medido por el promedio de los retrasos en el desplazamiento, haya aumentado sustancialmente en los últimos años. El costo promedio privado por hora de desplazamiento en automóvil ha disminuido con las mejoras sustanciales en la comodidad del automóvil, pero no lo suficiente para compensar por completo el aumento en el retraso.

La recomendación habitual de los economistas para hacer frente a las externalidades negativas es, como sabemos, gravar la actividad que las produce. Recientemente, Londres dio un gran paso en esa dirección al imponer una tarifa considerable por conducir en el centro de la ciudad entre semana. El resultado ha sido impresionante: una reducción de 20% en el tráfico de vehículos en Londres. Pero es incierto que este éxito se pueda duplicar en los Estados Unidos. Antes de la imposición de la tarifa de desplazamiento en Londres, 85% de las personas que viajaban a diario de su casa al trabajo ya hacían uso de autobuses y otras formas de transporte público en lugar de viajar

[22] Véase 42 U.S.C., 7651-7651 *o*; 58 Fed. Reg. 15634 (1993); Madison Gas & Electric Co. *vs*. EPA, 4 F.3d, 529 (7th Cir., 1993), 25 F.3D, 526 (7th Cir., 1994).

en automóvil. Esto indicaba que la mayoría de los viajeros pensaban que ese transporte era una buena alternativa a la conducción y, lo que es más importante, que el sistema de transporte público podía absorber perfectamente a los viajeros adicionales. Una disminución de 20% en el desplazamiento en automóvil se traduce en una migración de sólo 3% hacia el transporte público si los viajeros en automóvil son sólo 15% del número total de viajeros antes de que se imponga la tarifa. Es cierto que hay otros métodos para ahorrar en la conducción además de cambiar a otro modo de transporte, como compartir el automóvil. Pero compartir el automóvil tiene las características que no les agradan a las personas a quienes no les gusta el transporte público: menos privacidad y flexibilidad que conducir solo.

La resistencia a una tarifa de desplazamiento sería grande en las ciudades donde no se tienen buenas alternativas de transporte público, que son la mayoría de las ciudades de los Estados Unidos. (En promedio, sólo 2% de los viajeros estadunidenses usan el transporte público.) En tales ciudades, las tarifas altas de desplazamiento reducirían el número de viajeros, perjudicando a los negocios de la ciudad. Y, al reducir el congestionamiento y, por lo tanto, el costo del desplazamiento en automóvil, una tarifa de desplazamiento rígida podría tener sólo un efecto menor sobre el congestionamiento. De hecho, la tarifa inducirá a algunos viajeros a sustituir el transporte público, si tienen un costo de tiempo alto, por la conducción. Una sugerencia frecuente para combatir el congestionamiento del tráfico son las horas de trabajo escalonadas. Una sugerencia favorita de los economistas que tendría un efecto similar sería hacer que la tarifa de desplazamiento varíe según la hora del día, de modo que sería más alta durante las horas pico; pero estas sugerencias involucran un costo oculto: todo lo que reduce el congestionamiento atrae más tráfico al reducir los costos de tiempo de conducción y, entonces, el tráfico adicional crea, o vuelve a crear, los costos de tiempo.

XIV.8. Depósitos por recipientes obligatorios

Resulta difícil controlar la contaminación estética creada al tirar latas de cerveza y de refrescos, y botellas, a lo largo de las carreteras y en los parques. El costo de recolectar la basura es elevado y el costo de aprehender a quienes la tiran también es elevado. Las multas por tirar basura son cuantiosas en reconocimiento por la baja probabilidad de la aprehensión, pero tendrían que ser mucho mayores para tener un efecto disuasivo considerable… y en el capítulo VII vimos que los castigos severos pueden ser muy costosos en términos sociales. El instrumento del depósito obligatorio por recipientes, adoptado en varios estados, ha parecido una respuesta seductora. El distribuidor detallista debe hacer que sus clientes depositen una moneda de 5 o de 10 centavos de dólar por cada recipiente que compren, pero debe devolver el depó-

sito si el cliente devuelve el recipiente. El hecho de que, en principio, cada centavo del "impuesto" recolectado deba ser devuelto al contribuyente es un buen ejemplo de la diferencia entre usar los impuestos para elevar la recaudación y para regular el comportamiento. El depósito obligatorio vuelve costoso para el consumidor tirar basura, ya que pierde su depósito si no devuelve el recipiente, pero no genera ningún ingreso para nadie, ya que la competencia impedirá que los distribuidores detallistas se beneficien de los depósitos que retienen.

El problema es que no se hace ningún esfuerzo para estimar el costo de tirar basura y relacionarlo con el nivel del impuesto (depósito). Esta omisión podría deberse al hecho de que quienes proponen el impuesto lo consideran sin costo porque el cliente sólo tiene que devolver el recipiente para que le regresen su dinero. Su razonamiento olvida los costos no pecuniarios. El cliente incurre en un costo de almacenamiento y en un costo de tiempo al devolver estos recipientes de bebidas en lugar de tirarlos. Si el impuesto se fija a un nivel menor que su costo, la gente no devolverá sus recipientes y el único efecto del plan será el de aumentar los costos contables de los distribuidores detallistas. Sus ganancias no aumentarán incluso si ninguno de los clientes devolviera los recipientes y en consecuencia el detallista no tuviera que regresar nunca el depósito; el comercio minorista es un sector competitivo, por lo que cualquier rendimiento inesperado derivado del cobro del depósito del contenedor se verá eliminado por la reducción en los precios de los productos.

Si el "impuesto" de depósito se fija a un nivel suficientemente elevado para inducir las devoluciones, impondrá un costo social medido por el tiempo y otros costos de los consumidores que devuelvan los recipientes (*cf.* antes la gráfica XIV.3). Si ese costo es igual al costo social de la basura que se previene, como ocurrirá si el depósito se fija a un nivel igual al de la reducción de los costos sociales de la basura al devolver una botella más, todo estaría bien; pero, si es mayor, entonces la regulación mediante la tributación del depósito obligatorio será ineficiente.

XIV.9. El análisis costo-beneficio como una herramienta reguladora

La herramienta básica de la regulación llevada a cabo de acuerdo con los principios económicos es el análisis costo-beneficio; y aquí haremos una pausa para considerar los puntos fuertes y débiles de esta herramienta.

El término "análisis costo-beneficio" tiene diversos significados y usos. Al nivel de generalidad mayor, es virtualmente sinónimo del uso normativo de la economía. En el otro extremo de la escala de generalidad, el término denota el uso del criterio de la maximización de la riqueza para evaluar los proyectos gubernamentales, como la construcción de una presa o la adquisición de un

sistema armamentista; las subvenciones y donativos gubernamentales, como los donativos para la investigación médica; y algo particularmente importante en este capítulo: las regulaciones gubernamentales, incluidas no sólo las regulaciones administrativas que se ocupan de la salud, el ambiente y otras actividades fuertemente reguladas, sino también las leyes reglamentarias y las doctrinas y los fallos del derecho común. Pero el análisis costo-beneficio podría tener una utilidad positiva tanto como normativa al explicar o pronosticar las decisiones gubernamentales que se encuentran relativamente aisladas de la operación de la política del grupo de interés; un punto ilustrado por la teoría positiva del derecho común que se presentó en la segunda parte de este libro.

A lo largo de un eje de definición diferente, el análisis costo-beneficio puede referirse a: un método de evaluación pura, realizado sin tomar en cuenta el posible uso de sus resultados en una decisión; un insumo de la decisión, pues el tomador de decisiones está en libertad para rechazar los resultados del análisis sobre la base de otras consideraciones; o el exclusivo método de la decisión. Cuando el análisis costo-beneficio es sólo un insumo de la decisión, no hay necesidad de insistir en su adecuación como un principio normativo, siempre que se acepte la riqueza como un valor social, aunque no el único. Así empleado, el análisis costo-beneficio sólo obliga al tomador de decisiones a confrontar los costos de un curso de acción que se proponga.[23]

Las limitaciones cognitivas humanas proveen un argumento adicional en favor del uso del análisis costo-beneficio para la regulación y otros propósitos públicos. Contra la opinión de que el análisis económico desprecia las variables "suaves" (no cuantificables), de hecho el pensamiento ordinario está sesgado en contra de dar a las variables *cuantitativas* el peso debido. Cuando las dependencias reguladoras federales aplican el análisis costo-beneficio para regular los riesgos para la salud y la seguridad, se descubren, como señalamos anteriormente, estimaciones anómalas del valor de la vida que no se defenderían si se conocieran generalmente. Desde luego, si las limitaciones cognitivas afectan el comportamiento del mercado, los precios en los que se basa el análisis costo-beneficio no serán una herramienta totalmente confiable para disciplinar el pensamiento. Pero, por lo menos cuando las regulaciones se basen en el análisis costo-beneficio, quienes deseen cuestionarlas en los tribunales podrían señalar en cuáles sentidos fueron irrazonables los supuestos empleados en dicho análisis. Resulta más difícil cuestionar una regulación que se basa por entero en consideraciones cualitativas, como la "equidad".

[23] "El análisis costo-beneficio se intentó desde el principio como una estrategia para limitar la intervención de la política en las decisiones de la inversión pública." Theodore M. Porter, *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life,* 189 (1995), al discutir el uso del análisis costo-beneficio por el Cuerpo de Ingenieros del Ejército Estadunidense, que empezó a utilizar tal análisis en las propuestas de presas, bahías y otras obras públicas en el decenio de 1920.

Hablando de limitaciones cognitivas humanas, si los valores de las propiedades se derrumbaran a causa de un temor irracional al contagio de los pacientes en un hospital, ¿debiera eso contar como un costo en la decisión de dónde ubicar el hospital? Los costos de los temores irracionales para quienes los tienen son costos reales. Pero, si se ignoran y se construye el hospital, y los vecinos no se infectan, tenderán a disiparse los temores de contagio, de modo que los costos, aunque son reales, sólo lo son a corto plazo. Así, los temores irracionales debieran ponderarse sólo si es improbable que se despejen algún día, como en el caso del miedo a las alturas (seguras), que es resultado de nuestros genes más que de la ignorancia.

Un problema difícil en la aplicación del análisis costo-beneficio a los problemas ambientales es el de la valuación de los valores ambientales intangibles. Preguntar a la gente cuánto pagaría por salvar a un miembro de una especie en peligro de extinción no es satisfactorio, porque no *compramos* especies en peligro, de modo que mientras que al preguntar a una persona cuánto pagaría por una pasta de dientes obtendremos una respuesta con sentido, al preguntarle cuánto vale para ella una especie en peligro no obtendremos tal respuesta. Otro problema con estas "evaluaciones contingentes", como se las llama, es que es probable que la respuesta dependa de cuántas contingencias se le pide que valore al encuestado. Cuanto más se le pregunte, menor será la valoración de cada una, ya que tiene recursos limitados.

Una cuestión recurrente en el análisis costo-beneficio es la elección de la tasa de descuento,[24] algo que ya hemos encontrado en el contexto diferente de la valuación de la pérdida de ganancias futuras en los casos de perjuicios personales (véase antes VI.11). Supongamos una regulación ambiental cuya implementación costará 100 millones de dólares, todo ello en el primer año; se espera que esta regulación rinda beneficios monetarios de 10 millones de dólares anuales durante los próximos 100 años. ¿Cuál tasa debiera utilizarse para descontar los beneficios futuros hasta un valor presente que pueda usarse para comparar con el costo presente de 100 millones de dólares? Deben considerarse tres posibilidades. Una consiste en usar una tasa de descuento de cero alegando que no tenemos derecho a valuar el bienestar de las generaciones futuras en menos de lo que valuamos nuestro propio bienestar. Pero, aun si se aceptara la premisa, no se sigue la conclusión, pues omite la posibilidad de invertir el dinero que necesitaríamos para implementar la regulación y usar los intereses en el financiamiento de un proyecto futuro. Por ejemplo, podrían usarse ahora 50 millones de dólares para generar 50 años de beneficios, e invertir los otros 50 millones de dólares para permitir que se obtengan los siguientes 50 años de beneficios con los 50 años de intereses de la inversión.

[24] Véase Edward R. Morrison, "Comment, Judicial Review of Discount Rates Used in Regulatory Cost-Benefit Analysis", 65 *U. Chi. L. Rev.,* 1333 (1998).

Una segunda posibilidad consiste en emplear la tasa de interés comercial, como en el ejemplo que acabamos de presentar, y la tercera consiste en usar una tasa de interés "social", es decir, una tasa de interés que refleje una decisión gubernamental sobre la preferencia socialmente correcta a favor del valor presente sobre el valor futuro. La tasa comercial refleja las preferencias de la generación actual: es la que está invirtiendo, y, si la tasa de interés no es suficientemente elevada para superar su preferencia por el consumo corriente, no invertirá. (Recuérdese que, entre mayor sea la tasa de interés usada para descontar los valores futuros a los valores presentes, menores serán los valores presentes.) Si el gobierno decide que la gente está siendo demasiado egoísta al decidir cuánto gastar ahora en beneficio de las generaciones futuras, podrá elegir una tasa de descuento para usarla en el análisis costo-beneficio de sus proyectos que sea menor a la tasa de interés comercial; esto ponderará automáticamente el futuro en mayor medida que el mercado. Pero, otra vez, una tasa de descuento elevada implica tasas de interés elevadas (los prestamistas potenciales demandarán una tasa de interés muy elevada a fin de compensarse por diferir el consumo si no valúan mucho el consumo futuro), lo que significa que, si el costo del proyecto se ahorra e invierte, el interés compuesto creará en el futuro un fondo grande para el financiamiento de tales proyectos. Véase más acerca de estas cuestiones en la siguiente sección de este capítulo.

XIV.10. Los límites del análisis costo-beneficio: el caso del calentamiento global

Un gas de efecto invernadero es un gas que tiende a bloquear el calor reflejado desde la superficie terrestre a la vez que permite que el calor del sol penetre hasta redoblar el calor de ésta; así pues, entre más elevada sea la concentración de tales gases en la atmósfera, más caliente se volverá la superficie.[25] El gas de efecto invernadero más importante es el dióxido de carbono, que se emite a la atmósfera por varias actividades humanas, principalmente la quema de combustibles fósiles (petróleo, carbón y gas natural) y la de bosques; esos combustibles, y por supuesto la madera, están hechos de carbono y al quemarlos se libera carbono en la forma de dióxido de carbono.

Las décadas recientes han registrado anualmente grandes incrementos (cada vez más intensos) en las emisiones de dióxido de carbono, a causa del crecimiento de la población y de la actividad económica. Los aumentos de las emisiones han sido acompañados por sustanciales aumentos en la concentración atmosférica del dióxido de carbono, así como por aumentos en la

[25] Véanse William D. Nordhaus, *Managing the Global Commons: The Economics of Climate Change* (1994); Richard A. Posner, *Catastrophe: Risk and Response*, 317-318 (2004) (referencias del índice al calentamiento global); Martin L. Weitzman, "A Review of *The Stern Review on the Economics of Climate Change*", 45 *J. Econ. Lit.*, 703 (2007).

temperatura a escala global. Actualmente hay un consenso científico en el sentido de que estas variaciones están ligadas: las emisiones incrementan la concentración atmosférica del dióxido de carbono, y el aumento de esta concentración provoca que se eleve la temperatura a nivel global.

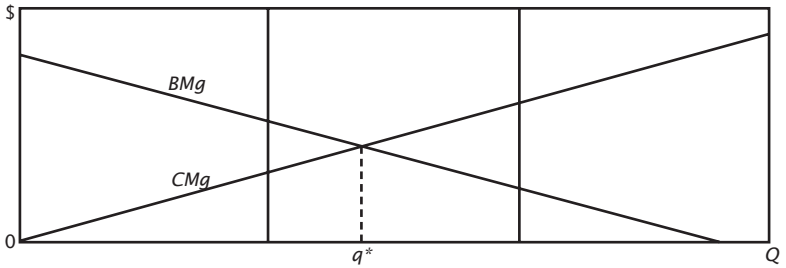
El Protocolo de Kioto, que los Estados Unidos se negaron a suscribir (además de que el dióxido de carbono no ha sido clasificado en este país como contaminante según la Ley del Aire Limpio), demanda que las naciones signatarias reduzcan sus emisiones de dióxido de carbono por debajo del nivel en que se encontraban hacia fines del decenio de 1990, presumiblemente al imponer impuestos a la energía eléctrica o a las emisiones, que induzcan al remplazo de actividades basadas en la quema de combustibles y que, por lo demás, fomenten un uso más económico de dichos combustibles. Sin embargo, el Protocolo exceptúa de tales medidas a los países en desarrollo, como China, responsable de emisiones de notable envergadura y constantemente crecientes, o como Brasil, que está destruyendo grandes extensiones de la selva amazónica con la quema. El efecto de las emisiones de dióxido de carbono sobre la concentración atmosférica de dicho gas es acumulativo, debido a que el dióxido abandona la atmósfera (al ser absorbido por los océanos) a mucha menor velocidad que al ingresar a ella, y, por consiguiente, se registra una concentración que habrá de seguir creciendo aun cuando se ponga un tope a la tasa anual de emisiones. Ubicado entre este fenómeno y las excepciones concedidas, el efecto directo del Protocolo de Kioto no puede evitar ser modesto; no obstante, el costo del impuesto o de otras medidas regulatorias requeridas para cumplir con ese acuerdo habrá de ser inmenso. Comoquiera que sea, ese costo les dará a las industrias energéticas, así como a sus clientes —como las aerolíneas y los fabricantes de vehículos automotores—, un poderoso incentivo para financiar la investigación y desarrollo (I&D) diseñada para crear sustitutos del carbono que sean tan económicos como limpios y, asimismo, instrumentos para atrapar las emisiones en su lugar de origen, antes de que ingresen a la atmósfera, o incluso para remover el dióxido de carbono de ésta. Por supuesto, este efecto indirecto del Protocolo se multiplicaría grandemente si los Estados Unidos firmaran el acuerdo y, con la finalidad de cumplirlo, fijaran impuestos contundentes a las emisiones.

Dado el efecto prácticamente irreversible que tienen las emisiones de gases con efecto de invernadero sobre su concentración atmosférica, estabilizarla en una fecha futura podría requerir recortes bastante más profundos de las emisiones que si se diera comienzo ahora mismo al proceso de estabilización. Hacer recortes menos hondos en la actualidad, como lo requiere el Protocolo de Kioto, equivale a comprar una opción para garantizar que en algún tiempo futuro el calentamiento global sea detenido o disminuya a un costo más bajo. En el caso de que la investigación encuentre que el problema del calentamiento global no es serio, dicha opción no sería elegida.

Para ilustrar el punto, supongamos que hay 70% de probabilidades de que en 2024 el calentamiento global sea causante de una pérdida social de un billón de dólares (al valor presente) y 30% de probabilidades de que no la provoque, y que la posible pérdida pueda ser evitada imponiendo hoy en día controles de emisión que le costarán a la sociedad 500 000 millones de dólares (para efectos de simplificación, demos por supuesto que el costo será cubierto en su integridad el primer año). Podría parecer que, puesto que la pérdida descontada por el calentamiento global en 2024 sea de 700 000 millones de dólares, la imposición de un control de emisiones en nuestros días tendría justificación desde el ángulo de los costos. Pero supongamos que en 2014 descubrimos de manera irrefutable que ha de sobrevenir un desenlace funesto (un billón) para 2024. Más aún, supóngase que si posponemos la imposición de un control de emisiones hasta 2014 todavía podremos evitar la pérdida de un billón. Entonces estaría claro que debiéramos aguardar, no solamente por la razón obvia de que el valor presente de 500 000 millones a ser gastados en 10 años es menor a 500 000 millones de dólares (a una tasa de descuento de 3% es aproximadamente de 425 000 millones), sino asimismo, lo que es más interesante, porque hay 30% de posibilidades de que no tengamos que incurrir en *ningún* costo por el control de emisiones. Como resultado, el costo esperado de los controles pospuestos no es de 425 000 millones, sino de sólo 70% de esa cantidad, esto es, 297 500 millones de dólares, lo que suma bastante menos que 500 000 millones de dólares. La diferencia radica en el valor de aguardar.

Pero ahora supongamos que el control de emisiones impuesto hoy en día —a fin de cumplir con el Protocolo de Kioto— que costara a la sociedad 100 000 millones, merced a una aceleración del ritmo de avance tecnológico, redujera de 500 000 millones a 250 000 millones el costo, en 2014, de alejar el espectro de la pérdida de un billón de dólares que, según se espera, causaría el calentamiento global en 2024. Después de descontar al valor presente 3% y 70% para reflejar la posibilidad de 30% de que ese control de emisiones no sea necesario, la cifra de 250 000 millones se encoge hasta 170 000 millones de dólares. Esto significa 127 500 millones de dólares menos que el enfoque superficialmente atractivo de "aguardemos a ver qué pasa" (297 500 millones menos 170 000 millones). Por supuesto, existe un precio a pagar por el enfoque modificado: 100 000 millones. Empero, el valor es más grande que el precio.

Puede ser que el Protocolo de Kioto constituya una respuesta demasiado limitada al calentamiento global si, en lugar de enfocarnos en el calentamiento global de carácter gradual, lo hacemos en el abrupto. Durante varios periodos a través de la historia de la Tierra, en el espacio de una década o de un lapso de tiempo menor han acontecido incrementos drásticos en la temperatura a escala global. Un aumento semejante que ocurriera hoy tendría efectos devastadores sobre la agricultura y sobre las ciudades costeras, e incluso podría causar un viraje en la Corriente del Golfo que tuviera como



GRÁFICA XIV.4. *El enfoque tolerable de las ventanas*

consecuencia proporcionarle a Europa un clima siberiano. Con todo, es imposible asignar una probabilidad numérica a tamaña catástrofe. He aquí un ejemplo del concepto de incertidumbre, concebido como algo diferente del riesgo calculable, introducido en el capítulo I. Este escenario suscita la pregunta de cómo es que podría uno conducir un análisis costo-beneficio sobre las medidas a tomar para ahuyentar un calentamiento global de naturaleza abrupta si los costos esperados de la catástrofe no pueden ser estimados por la razón de que las probabilidades de que ésta ocurra resultan indiscernibles, ignotas.

Una respuesta se describe en la gráfica XIV.4. Los beneficios marginales y los costos marginales de las medidas dirigidas a reducir o eliminar un riesgo catastrófico se muestran allí como funciones de las precauciones tomadas, donde la cantidad óptima de tales precauciones (q^*) se obtiene por la intersección de las dos funciones. Aun cuando la incertidumbre hace imposible determinar la óptima, podemos averiguar lo suficiente sobre los beneficios y costos para poder crear la "ventana" formada por las dos líneas verticales dentro del recuadro. Al lado izquierdo del marco de la ventana, los beneficios derivados de un esfuerzo suplementario por eliminar o prevenir la catástrofe exceden con toda comodidad a los costos, mientras que en el lado derecho lo contrario es cierto. Si nos mantenemos adheridos a la ventana, aunque no nos será posible saber si nuestras medidas son las óptimas, cuando menos tendremos alguna base para tener confianza en que ni son flagrantemente inadecuadas ni son flagrantemente excesivas.

Otro ajuste que debe realizarse para poder emprender el análisis costo-beneficio en condiciones de incertidumbre consiste en introducir el "análisis inverso de costo-beneficio". Análogo a los estimados de la probabilidad de extracciones de las primas de los seguros, este tipo de análisis implica dividir lo que está gastando el gobierno a fin de impedir que un riesgo catastrófico en particular cobre cuerpo por el costo social que tendría ese acontecimiento si en efecto tuviese lugar. El resultado es la probabilidad implícita de la catástrofe. Recuérdese que el costo esperado es el producto de la probabilidad y la consecuencia (la pérdida): $C = PL$. Si P y L son conocidas, se puede calcu-

lar *C*. Si en cambio *C* y *L* son conocidas, se puede calcular *P*: Si se han de gastar mil millones (*C*) para ahuyentar un desastre que, en caso de sobrevenir, impondrá una pérdida (*L*) de 100 000 millones de dólares, entonces $P = C/L = 0.01$.

Si, calculada de esta forma, *P* diverge agudamente de las estimaciones independientes al respecto, esto constituye un indicio de que la sociedad está gastando demasiado o demasiado poco en evitar *L*. Se trata sólo de un indicio, a causa de la distinción entre los costos marginales y totales y los beneficios. Supongamos que *P* no es 0.01 sino 0.1, de modo que el costo esperado de la catástrofe no es de 1 000 millones sino de 10 000 millones de dólares. De ahí no se sigue que debiéramos gastar 10 000 millones o, en efecto, nada más de un millón para impedir la catástrofe. Tal vez gastar sólo mil millones reduciría el costo esperado de la catástrofe de 10 000 millones hasta no más de 500 millones y ningún gasto adicional traería consigo una reducción adicional, o al menos una reducción adicional justificada por el costo. Por ejemplo, si gastar otros mil millones fuera a reducir el costo esperado de 500 millones de dólares a cero, ésa sería una mala inversión, al menos si se ignora la aversión al riesgo.

Supongamos que el calentamiento global abrupto y catastrófico provocara una reducción permanente de una quinta parte del PIB de los Estados Unidos, que en la actualidad es de 13 billones de dólares. Como se da por sentado que hay de manera permanente una pérdida anual de 2.6 billones de dólares, el valor presente de la pérdida total motivada por ese desastre, a una tasa de descuento de 3%, sería de alrededor de 85 billones de dólares. En los tiempos que corren el gobierno está gastando aproximadamente 2 000 millones en subsidiar la I&D sobre el calentamiento global. Si esa cantidad es la adecuada para evitar el escollo que implicaría el calentamiento global de carácter abrupto, la implicación es que la probabilidad de ocurrencia de semejante catástrofe es de solamente uno entre varios cientos de miles ($P = C/L =$ \$2 000 millones/\$85 billones = 0.000024). Ese cálculo parece demasiado bajo aun si no sabemos cuál sería el gasto óptimo.

Una alternativa a los subsidios vista como un medio para fomentar la creación de tecnología que redujera o eliminara el calentamiento global sería, tal como se mencionó antes, fijar una tributación a las emisiones, empleando así un instrumento que impeliera a lograr avances tecnológicos en esa dirección. La ventaja de semejante tributación sobre un subsidio residiría en su capacidad para mantener al gobierno alejado del negocio de elegir a los ganadores: las industrias afectadas habrían de decidir a qué clases de I&D se debía prestar apoyo. Asimismo, la mayor parte del impuesto podría ser parcialmente compensada reduciendo otros impuestos, toda vez que un gravamen a las emisiones recaudaría ingresos al mismo tiempo que induciría a mayores gastos en investigación y desarrollo.

¿Pero no serían necesarios los subsidios para las tecnologías que no contaran con ningún mercado, como aquellas destinadas a remover dióxido de

carbono de la atmósfera? No existiría una demanda privada debido a que, en contraste con las tecnologías que reducen las emisiones, las dedicadas a remover de la atmósfera dióxido de carbono ya emitido (el secuestro de carbono) no reducirían la carga fiscal de ningún emisor. Pero este problema puede resolverse fácilmente convirtiendo al impuesto en un impuesto sobre las emisiones *netas*. Entonces, una planta generadora de electricidad u otro emisor podrían reducir su carga fiscal removiendo dióxido de carbono de la atmósfera y a la vez reduciendo sus propias emisiones de éste en dirección a la atmósfera. Si lo removieran más de lo que lo emitieran, obtendrían un crédito fiscal.

¿Debería el impuesto aplicarse a las emisiones o a los combustibles fósiles (un impuesto a la gasolina o al contenido de BTU? Resulta más barato hacer cumplir un impuesto a la energía ya que no existe la necesidad de monitorear las emisiones. Pero sólo el impuesto a las emisiones sería efectivo para inducir el secuestro de carbono, porque éste reduce la cantidad de dióxido de carbono en la atmósfera sin recortar la demanda de combustibles fósiles. Un impuesto a la gasolina reduce su demanda, pero no promueve los esfuerzos para evitar que el dióxido de carbono emitido por la quema de la gasolina que sigue siendo producida ingrese a la atmósfera.

Al centrarnos en el riesgo del calentamiento global de tipo abrupto, podemos holgadamente evadir el embarazoso problema de elegir una tasa de descuento, cuestión que entraña una dificultad particularmente difícil de sortear cuando las preocupaciones giran en torno al calentamiento global de tipo gradual. Supongamos que un gasto de 10 000 millones de dólares para poner un límite o techo a las emisiones hoy en día no tendría efecto alguno en las condiciones de bienestar de la humanidad durante este siglo pero, al disminuir el ritmo del calentamiento, produciría un ahorro en los costos sociales de 100 000 millones de dólares en 2100. A una tasa de descuento de 3%, el valor presente de 100 000 millones a 100 años del año en curso es de solamente 5 000 millones de dólares. A la luz de ese dato, parecería que el gasto de 10 000 millones realizado el día de hoy sería una inversión muy mala. Podría esperarse que la misma cantidad de dinero invertida en instrumentos financieros creciera a 192 000 millones de dólares para 2100, suponiendo una tasa de interés real de 3% para los próximos 100 años (aunque, de hecho, las tasas de interés no pueden ser pronosticadas para un periodo tan prolongado). Si los fondos fueran entonces desembolsados en beneficio de las víctimas del calentamiento global, éstas se encontrarían en una mejor situación que si hubiera sido evitado el costo de 100 000 millones del calentamiento global en que, según se da por sentado, se incurriría en ese año. Además, inversiones más conservadoras arrojarían ingresos esperados más sustanciales: 10% o más en vez de 3 por ciento.

Con todo, invertir 10 000 millones en fondos para las futuras víctimas del calentamiento global no es una verdadera alternativa a gastarlos ahora mismo.

No habrá de crearse un fondo de esa índole. En estas circunstancias, descontar valores futuros a los valores presentes no sería un método para ayudar a la gente a decidir cómo administrar sus asuntos de la manera que conduzca con mayor certeza a la maximización de su bienestar.

En el mejor de los casos, una tasa de descuento basada en las tasas de interés del mercado tiende a destruir los intereses de las generaciones del futuro más remoto. A una tasa de descuento de 5%, el valor presente de un dólar a ser recibido en 100 años son sólo tres cuartos de un centavo; y, si sustituimos el dinero por vidas, entonces por salvar una vida este año debiéramos estar dispuestos a sacrificar 150 vidas de aquí a una centuria. Aun así, no descontar de ningún modo costos futuros sería absurdo, pues entonces el valor presente de los beneficios conferidos a nuestros descendientes remotos se acercaría al infinito. Las medidas adoptadas hoy para detener el calentamiento global otorgarían beneficios no sólo en 2100 sino en cada año subsecuente, quizás a lo largo de millones de años. El valor presente de 100 000 millones recibidos año con año durante un millón de años a una tasa de descuento de 0% es de 100 trillones de dólares, lo cual suma más de lo que Greenpeace anhela que se gaste en poner un coto a las emisiones de gases de efecto invernadero.

Una manera posible de manejar con sutileza el problema de cuánto peso ha de darse al bienestar de las generaciones del futuro remoto consiste en trasladar un valor presente descontado a uno no descontado recortando el horizonte de tiempo para la consideración de los costos y beneficios. Por ejemplo, el valor presente de un flujo infinito de costos descontados a 4% cada año es igual a la suma descontada de esos costos a lo largo de 25 años, en tanto que el valor presente de un flujo infinito de costos descontados a 1% por un año es igual a la suma no descontada de esos costos durante 100 años. La fórmula para el valor presente de 1 dólar por año para siempre es de $\$1/r$, donde r es la tasa de descuento. De este modo, si r es 4%, el valor presente es de 25 dólares, y esto es igual a un flujo no descontado de 1 dólar por año durante 25 años. Si r es 1%, el equivalente no descontado son 100 años.

Un modo de abogar por la tasa de 4% (es decir, por truncar nuestra preocupación por el bienestar futuro dejándolo en 25 años) consiste en decir que estamos dispuestos a ponderar el bienestar de la siguiente generación tan concienzudamente como nuestro propio bienestar pero que hasta ahí para nuestro pendiente por el futuro. Una forma de argumentar a favor de la tasa de 1% consiste en decir que estamos dispuestos a dar la misma ponderación al bienestar de todos y cada uno de los que viven en este siglo, lo que nos incluye a nosotros, a nuestros hijos y a nuestros nietos, pero más allá de ese punto no nos importa.

Una forma más de enfocar las consecuencias futuras sin recurrir al descuento convencional consiste en agregar riesgos a lo largo del tiempo en lugar de expresarlos en términos anualizados. Si lo que nos concierne es aquello que puede ocurrir en el decurso del siguiente siglo, entonces, en lugar de preguntar

cuál es la probabilidad anual del calentamiento global de tipo abrupto, podríamos preguntar cuál es la probabilidad de que tal acontecimiento suceda en los próximos 100 años. Una probabilidad anual de uno en un millón se traduce en una probabilidad por centuria de casi 1 en 10 000. Esto puede ser lo suficientemente elevado —considerando las repercusiones de que tenga lugar ese acontecimiento— para justificar el gasto de muchos miles de millones de dólares a fin de evitarlo.

XIV.11. TELEVISIÓN POR CABLE: PROBLEMAS DE *COPYRIGHT* Y DEL MONOPOLIO LOCAL

Un sistema de televisión por cable obtiene sus señales de una antena maestra que las toma de estaciones de transmisión cercanas o de una repetición por microondas o de un sistema satelital que las acerca desde una antena distante que recibe las señales de estaciones de transmisión cercanas. Al traer señales distantes, el operador del cable aumenta el número de señales disponibles en el mercado local en el que opera. A las estaciones de difusión locales no les gusta esto porque pierden televidentes a manos de la señal nueva y con ello ingresos de publicidad. A los propietarios de los *copyrights* sobre los programas que se pasan en la señal distante tampoco les gusta la importación (gratuita). Al aumentar la audiencia para el programa transmitido por la estación de origen, la importación aumenta las regalías que el propietario del *copyright* puede extraer de esa estación, pero esta ganancia se ve contrarrestada por la pérdida de regalías de la estación que transmite el programa en competencia con el sistema de cable. Si los efectos se contrarrestaran perfectamente, el propietario del *copyright* sería indiferente al hecho de que el operador del cable importara señales distantes. Pero la importación impide que el propietario del *copyright* ofrezca derechos exclusivos a una estación que compita con el sistema de cable, y los derechos de programas exclusivos tienen una prima porque permiten que una estación diferencie su programación de las programaciones de sus competidores. Dicho de otro modo, los derechos de transmisión no exclusivos pueden crear una externalidad de congestionamiento (véase antes III.1, y especialmente la discusión del derecho al uso exclusivo del nombre y la imagen propios en la publicidad en XI.4).

Hay dos métodos para regular señales distantes: los derechos de propiedad intelectual o de autor (*copyright*) y la regulación federal de las operaciones de televisión por cable (por la Comisión Federal de Comunicaciones, FCC). El último método es, lógicamente, el tema principal de este capítulo. La FCC adoptó limitaciones sobre la importación. La Corte Suprema sostuvo que la importación no es una infracción del *copyright*.[26] La corte razonó que la

[26] Fortnightly Corp. *vs.* United Artists Television, Inc., 392 U.S., 390 (1968). Véase asimismo Teleprompter Corp. *vs.* CBS, Inc., 415 U.S., 394 (1974).

construcción de un sistema de televisión por cable es como la erección de una antena muy alta por un televidente individual. La analogía no ilumina el problema económico, que es el dilema entre los beneficios sociales del aumento de la producción de obras con *copyright* que se genera otorgando mayor protección al propietario del *copyright* y los costos sociales impuestos por el aumento del costo marginal del uso de obras con *copyright* (véase antes XI.5, y adviértase la analogía con el precio óptimo de la producción de un monopolista natural). La consideración, por parte de la FCC, de la economía de las señales distantes estaba sesgada por el hecho de que quienes comparecieron ante ella incluyeron a las estaciones de televisión locales y a los propietarios del *copyright*. La competencia no es un cuasidelito del derecho común, pero una dependencia reguladora podría otorgar una consideración amable al interés de una empresa por liberarse de la competencia, y a menudo lo hace. Se creía que la carga principal de la competencia por cable sería soportada por las estaciones independientes, principalmente las que usan bandas de UHF del espectro electromagnético. Dado que la Comisión, en sus políticas de asignación y licenciamiento, había apoyado asiduamente el crecimiento de la televisión de UHF durante muchos años, se inclinaba naturalmente a simpatizar con las quejas de los UHF independientes por la competencia de la televisión por cable.[27]

Otro problema regulatorio en este ámbito surge por el hecho de que una compañía que desee proveer el servicio de televisión por cable deberá obtener una franquicia de las autoridades municipales. Sus tarifas para los suscriptores podrían estar sujetas también a la regulación de la comisión estatal de servicios públicos.[28] En virtud de que la distribución por alambrado de las señales de televisión al hogar se asemeja en sus aspectos técnicos al servicio local de gas, agua y electricidad —los ejemplos tradicionales del monopolio natural—, no resulta sorprendente que haya un apoyo generalizado para la regulación de las tarifas de los suscriptores. Pero un enfoque alternativo, basado en el derecho de los contratos, podría ser superior.

Muchas compañías pueden construir y operar un sistema de televisión por cable en cualquier parte del país. Cuando no hubiera limitaciones a la entrada, esperaríamos que varias compañías se disputaran el privilegio de servir a cada comunidad donde haya una demanda considerable de televisión por cable. La compañía que ofreciera el mejor paquete de precios y ser-

[27] El Congreso ha hecho enmiendas a la ley del *copyright* a fin de extender su protección a las señales más distantes. La enmienda dispone que se otorgue licenciamiento obligatorio a las tarifas de regalías establecidas en la ley. Copyright Revision Act of 1976, 17 U.S.C., 111. Véase Stanley M. Besen, Willard G. Manning, Jr. y Bridger M. Mitchell, "Copyright Liability for Cable Television: Compulsory Licensing and the Coase Theorem", 21 *J. Law & Econ.*, 67 (1978). También han sido abrogadas las reglas de la FCC sobre las señales distantes.

[28] No obstante, la regulación de la franquicia y la regulación de la tarifa están sujetas a los límites de la legislación federal.

vicio suscribiría a la mayoría de los clientes, y si el servicio de televisión por cable local es realmente un monopolio natural, esa compañía tendría costos medios menores que sus competidores y los expulsaría del mercado.

En este punto, la empresa exitosa tendría un monopolio. ¿Estarían por lo tanto los suscriptores a su merced? No, si el nivel prometido de precio y servicio se hubiese especificado en un contrato con cada uno de los suscriptores potenciales, y la competencia entre los aspirantes rivales al mercado local debiera asegurar que los suscriptores potenciales recibieran un contrato vinculante. Ni el proceso de determinación de la empresa que prevaleciera impondría costos sustanciales bajo la forma de instalaciones duplicadas que resultaran redundantes. Ninguna empresa empezaría a construir su red de cables hasta después de un periodo de búsqueda de suscriptores, y, si ese periodo revelara una fuerte preferencia de los consumidores por una empresa, las otras le venderían a ésta sus contratos de suscripción. Si el periodo de oferta a posibles suscriptores no revelara ninguna preferencia fuerte por una empresa, y cada una de ellas tuviera suscriptores dispersos por todo el mercado, presumiblemente las empresas intercambiarían suscriptores hasta que cada una de ellas tuviera un área de mercado compacta por servir.

Dado que los costos de la contratación podrían ser muy elevados, un enfoque puramente contractual podría no ser satisfactorio. Habría dificultad para hacer a un lado al proveedor exitoso, así que un consumidor necesitaría un contrato a largo plazo para evitar ser explotado; no obstante, sería impracticable especificar y hacer cumplir el precio y otros términos del contrato por un tiempo prolongado[29] en lo que respecta a los intereses en juego, que para cada suscriptor serían de poca monta. Pero hay un enfoque regulador que preserva los elementos esenciales del enfoque contractual. El municipio, actuando como el agente contratante de sus residentes, podría solicitar pujas de solicitantes rivales para la franquicia (o las franquicias) de televisión por cable. El municipio compararía estas pujas y celebraría un contrato en firme con el solicitante que prometiera la mejor combinación de precio bajo y servicio generoso. Si le resultara difícil escoger entre diferentes paquetes de precio y servicio, podría realizar una encuesta de preferencias entre los residentes después de una campaña realizada por empresas solicitantes.

El enfoque contractual modificado podría parecer muy semejante al sistema actual de la franquicia del cable, porque el que recibe la franquicia se

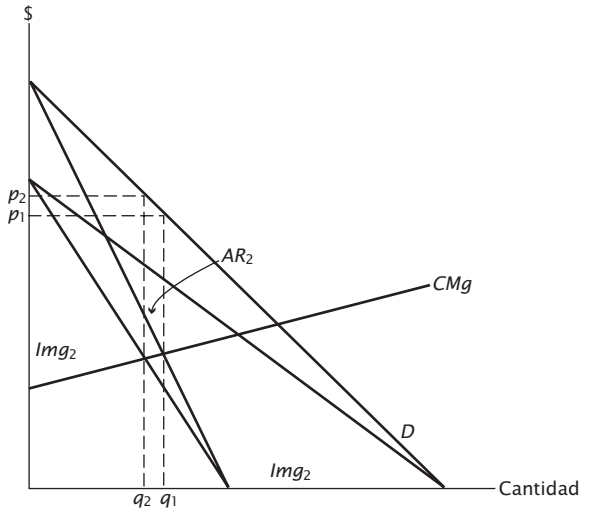
[29] Acerca de la economía de los contratos a largo plazo, véanse, por ejemplo, Keith J. Crocker y Scott E. Masten, "Pretia ex Machina? Prices and Process in Long-Term Contracts", 34 *J. Law and Econ.,* 69 (1991); Victor P. Goldberg y John R. Erickson, "Quantity and Price Adjustment in Long-Term Contracts: A Case Study of Petroleum Coke", 30 *J. Law & Econ.,* 369 (1987); J. Harold Mulherin, "Complexity in Long-Term Contracts: An Analysis of Natural Gas Provisions", 2 *J. Law, Econ. & Organization,* 105 (1986); Howard A. Shelanski y Peter G. Klein, "Empirical Research in Transaction Cost Economics: A Review and Assessment", 11 *J. Law, Econ. & Organization,* 335 (1995).

escoge de ordinario entre varios solicitantes. La diferencia es que, bajo el sistema actual, el municipio no se fija de ordinario como meta la obtención del mejor contrato posible para los suscriptores. Más bien, se trata de extraer concesiones de quien recibe la franquicia como el precio del permiso para cobrar un precio monopólico. Una franquicia de cable típica establecerá que el poseedor de la franquicia debe remitir al municipio un porcentaje de sus ingresos brutos y dedicar también cierto número de canales, sin cargo alguno, para servicios municipales tales como el de educación. Los residentes ganan algo de la franquicia como contribuyentes, pero pierden como consumidores. Pagan una tarifa de suscripción monopólica que es mayor, debido a las concesiones, de lo que sería si la compañía de televisión por cable fuera un monopolista sin restricciones, ya no digamos una compañía contractualmente obligada a proveer el servicio a un precio igual al costo. Esto se ilustra en la gráfica XIV.5, que compara el precio monopólico con un impuesto a los ingresos brutos y sin él.[30] El ejemplo muestra la diferencia existente entre la tributación por la regulación y la maximización del bienestar del consumidor por el contrato.[31]

[30] Para entender el efecto del impuesto, téngase en mente que desde el ángulo del vendedor la tabla de demanda es una tabla del ingreso medio —del precio recibido en vez del precio cobrado— en varios niveles de la producción. El impuesto altera la tabla de demanda de la compañía (el ingreso medio), reduciéndola por un porcentaje uniforme a cada nivel de la producción, según se muestra en la gráfica XIV.5. La tabla de ingreso marginal de la compañía también queda alterada, como se representa en el diagrama, mientras que la producción maximizadora de la ganancia resultante es más baja y el precio (incluido el impuesto) más elevado que antes de que fuera impuesta la tributación.

Supóngase que el municipio de nuestro ejemplo, en vez de fijar un impuesto a los ingresos brutos, cobra un impuesto plano por unidad de producción. Esto haría que, al observar que la curva de costo marginal se elevaba a una distancia uniforme, la empresa incrementara su precio y redujera su producción. Véase XVIII.3 después. Si el municipio sencillamente pone en subasta la franquicia de la televisión por cable buscando al postor que ofrezca la suma al contado más jugosa, éste se vería librado de cualesquiera controles sobre este precio o esta producción, de modo que establecería los mismos niveles de precio y de producción que los que impondría un monopolista no regulado. Durante el proceso de puja, cada aspirante se haría una idea del valor presente de las ganancias que podría obtener fijando su precio y su producción al nivel del monopolio, y luego ofrecería algo menor a dicho valor presente en la subasta. El pujador con el ofrecimiento máximo ganaría la subasta y, a continuación, procedería a establecer los niveles de precio y producción que había calculado que maximizarían sus ganancias monopólicas. Si la franquicia fuese exclusiva, los pujadores estarían dispuestos a apostar una cantidad mayor que si no fuera exclusiva, dado que en el primer caso el descuento que aplicarían a las ganancias monopólicas esperadas —tomando en cuenta la posibilidad de que otra empresa entrara al mercado y les arrebatara el triunfo en la competencia— sería más pequeño. Los municipios que adoptaran el método del pago al contado en la subasta podrían por ello sentirse tentados a otorgar franquicias exclusivas a fin de aumentar sus ingresos derivados del proceso subastador; pero, al mismo tiempo, estarían incrementando los costos del monopolio para sus residentes.

[31] Para una crítica del enfoque propuesto arriba, en la que se hace hincapié en los costos de transacción, véase Oliver E. Williamson, "Franchise Bidding for Natural Monopolies—In General and with Respect to CATV", 7 *Bell J. Econ. & Mgmt. Sci.,* 73 (1976).



GRÁFICA XIV.5

Si la contratación del derecho común (quizá con cierta regulación adicional) es una respuesta posible a la regulación de la televisión por cable, ¿por qué no es una respuesta posible para la regulación de otros monopolios naturales y, por lo tanto, una alternativa para el elaborado sistema regulador discutido en el capítulo anterior? En efecto, la regulación de las franquicias era el sistema de regulación empleado en los primeros días de la distribución eléctrica, los tranvías y los servicios telefónicos, todos ellos servicios del tipo de red como la televisión por cable. Pero también fue un tremendo fracaso, sobre todo, al parecer, porque las autoridades que otorgaban la franquicia no representaban en efecto el interés del consumidor y, por lo tanto, no acertaban a desempeñar el papel del comprador en el contrato convencional.

Además de este problema político, hay un problema económico más puro que nos hace regresar a la discusión de los contratos entre terratenientes e inquilinos en el capítulo III (véase antes III.11). La franquicia debe tener una duración fija, porque resulta imposible fijar términos sensatos (precio, servicio, etc.) para el futuro indefinido. Supongamos que la duración óptima es mucho más breve que la vida útil de muchos de los activos de quien obtiene la franquicia (derechos de vía, etc.). A la terminación del plazo de la franquicia habrá un problema de monopolio bilateral de cierta gravedad. El costo de la remoción de los activos será muy elevado para el antiguo poseedor de la franquicia (en particular porque pueden tener escaso o ningún valor en otros usos), pero el costo de reposición de los activos podría ser también muy elevado para la autoridad que otorgó la franquicia. Una solución simple sería que los activos pasaran a la propiedad de este último al terminar la franquicia. Pero

entonces, al igual que en el caso de las mejoras realizadas por el inquilino que pasaban a las manos del terrateniente al terminar el contrato, el que consiguió la franquicia invertiría deficientemente. Una alternativa mejor, que se encuentra en muchas ordenanzas de la televisión por cable, consiste en establecer, al momento de otorgar la franquicia, una fórmula para determinar el precio de los activos de la empresa que obtuvo la franquicia cuando ésta termine.[32]

XIV.12. La venta de las autopistas

Paralelamente al enfoque del contrato para regular la televisión por cable, ha surgido cada vez con mayor fuerza un movimiento dirigido a privatizar los caminos y carreteras de cuota o peaje. El estado de Indiana arrendó recientemente el Camino de Cuota de Indiana —una carretera de 252.5 kilómetros al norte de Indiana que conecta a Illinois con Ohio— a un consorcio español-australiano por un periodo de 75 años a cambio de 3 800 millones de dólares, pagados al contado.

La idea de privatizar los caminos de cuota es atractiva desde el punto de vista económico. Las empresas privadas tienen a la vez un poderoso incentivo financiero y una presión competitiva para esforzarse en minimizar sus costos —factores que operan débilmente o de manera nula en el caso de las dependencias públicas—, y sus decisiones en materia de fijación de precios y compras, incluidas las decisiones tocantes a los salarios y las relaciones laborales, no sufren distorsiones a causa de presiones políticas o de corrupción, siempre y cuando la sociedad cuente con instituciones efectivas para minimizar la interferencia política (engendradora de corrupción) en los mercados privados. No es que estas condiciones se puedan alcanzar con la mayor facilidad o sin costo alguno. La alta incidencia de determinación dudosa de los precios en la construcción y el mantenimiento de los caminos y carreteras es atribuible en parte a reglas de puja que, al intentar prevenir la corrupción, abren el paso a licitaciones amañadas. Por dar un ejemplo, si para prevenir la corrupción los contratos se otorgan siempre al pujador más bajo, habrá siempre una conspiración lista para arreglar la puja que, habiendo asignado a un postor para hacer la puja más baja, descubra quién de entre sus miem-

[32] La regulación se analiza después en XVI.11-XVI.14 (regulación de los mercados de valores y de la banca). Pero aquí cabe advertir que a veces se ha pensado que el tipo de análisis presentado en el texto justifica la regulación de los contratos de franquicia ordinaria (por ejemplo, las franquicias de comida rápida), una forma de regulación presente en muchos estados. Se ha encontrado que estas regulaciones son ineficientes. James A. Brickley, Frederick H. Dark y Michael S. Weisbach, "The Economic Effects of Franchise Termination Laws", 34 *J. Law & Econ.*, 101 (1991). Es difícil que esto cause sorpresa. No hay ningún problema de números abultados que impidiera una solución contractual directa. Las incertidumbres inherentes a un contrato a largo plazo con costos hundidos están presentes, pero no existe razón alguna para creer que esas dudas se resuelvan más fácilmente recurriendo al gobierno que a las partes contratantes.

bros la está engañando y se ha quedado con el contrato. En los casos donde el engaño en el seno de una agrupación de conspiradores es fácilmente detectable, es menos probable que éste fructifique y, en cambio, más seguro que la conspiración sea eficaz.

Comoquiera que sea, la privatización de los caminos y carreteras de cuota genera un problema de monopolio. Los automovilistas que no tienen buenas alternativas al uso de esas carreteras pueden ser obligados a pagar un peaje que sobrepasa con mucho los costos que, en términos de deterioro físico, efectos de congestionamiento, repercusiones sociales de la contaminación, etc., representan esos caminos, y habrá entre ellos quienes, estando indispuestos a pagar, busquen vías sustitutas que resultan ineficientes. Esta cuestión debiera llevarnos a preguntar qué era exactamente lo que vendía Indiana cuando arrendó el camino de cuota por 3 800 millones de dólares. Entre más altas sean las cuotas y mayor sea el margen de libertad del arrendatario para aumentarlas en el futuro, más elevado será el precio que el estado pueda fijar para el arriendo. Si el estado no impusiera limitaciones a las cuotas, estaría vendiendo un monopolio no regulado. Si el arriendo pudiera constreñir al arrendatario a cobrar cuotas exactamente iguales al costo de operar el camino de peaje (incluidos el mantenimiento, las reparaciones, la remoción de la nieve, el alumbrado y la recaudación de las cuotas), el precio de mercado del arrendamiento sería significativamente más bajo. En la medida en que el estado pretenda maximizar el beneficio obtenido del arrendamiento, estará creando una ineficiencia de asignación al conferir un poder monopólico al arrendatario, tal como vimos en el caso de las franquicias de televisión por cable.

Es difícil determinar si el precio de 3 800 millones de dólares por el Camino de Cuota de Indiana constituye el precio competitivo o el precio monopólico. Por un lado, se le prohíbe al arrendatario elevar las cuotas hasta 2016 (para los autos particulares: el peaje para los camiones se incrementó en 2010), e incluso a partir de esa fecha sólo lo podrá hacer cuando mucho por 2% al año, la tasa de inflación del consumidor o el incremento anual del PIB. Por otro lado, las cuotas aumentaron significativamente un poco antes del arrendamiento, y la autorización concedida al operador en 2010 para empezar a aumentar dichas cuotas anualmente en correspondencia con el aumento del PIB puede traerle ganancias súbitas, ya que el costo de operar esos caminos no podrá incrementarse a una tasa tan alta.

Con todo, la preocupación por el monopolio no viene a ser una objeción contundente al arrendamiento del camino de cuota, puesto que todos los impuestos viables provocan efectos similarmente monopólicos en el sentido de que abren una cuña entre el costo y el precio. (Esta calificación es asimismo aplicable a la crítica de las cuotas de franquicia elevadas en la televisión por cable.) Supongamos que el precio de arrendamiento habría sido de sólo 2 mil millones de dólares si el estado hubiera impuesto limitaciones más rigurosas al aumento en las cuotas. Entonces el estado recibiría 1 800 millones

menos en ingresos y es de presumir que cubriría la diferencia elevando las tasas de impuestos o fijando impuestos adicionales, y estas medidas habrían tenido efectos de asignación similares a los de las cuotas más altas cobradas por el arrendatario del camino de peaje. Si el asunto del monopolio es considerado por lo tanto como tablas o un juego empatado, el efecto principal del arrendamiento será de signo positivo por consistir en la reducción del costo ajustado a la calidad que conlleve el funcionamiento de la carretera de cuota.

Los opositores argumentan que la privatización, por lo menos cuando cobra la forma de una venta o de un arrendamiento a largo plazo de la propiedad gubernamental por un pago al contado, se traduce en la futura miseria en cuanto que priva al gobierno de un activo que produce ingresos. Es verdad, empero, que Indiana no se está desprendiendo de un activo, sino simplemente haciendo que mude de apariencia: pasa de ser una carretera a efectivo. Pero, dado que a los funcionarios electos se les presentan horizontes truncados, existe el peligro de que el estado no invierta el efectivo juiciosamente para un largo plazo, sino que lo dilapide en proyectos a corto plazo. En efecto, Grecia fue acusada de haber vendido activos nacionales de carácter lucrativo por dinero en efectivo que gastó en lugar de invertirlo, y, a resultas de ello, se ha disminuido el ingreso requerido para satisfacer sus colosales obligaciones financieras. Si nos acercamos a casa, los fondos de pensión pública se hallan notoriamente subfinanciados debido a que los funcionarios proyectan de manera poco realista tener beneficios elevados con base en la inversión, lo que les permite predecir que las pequeñas inversiones anuales crecerán con rapidez.

Un último punto es que resulta menos probable que el congestionamiento ocurra en los caminos y carreteras de cuota de propiedad privada que en los de propiedad pública. La entidad pública será más receptiva ante quejas de los automovilistas que prefieren un embotellamiento a cuotas onerosas, ya que son votantes y superan en número a los conductores cuyo alto costo de tiempo (con la implicación de que reciben ingresos más altos) hace que estén dispuestos a pagar una prima porque circulen menos vehículos.

Lecturas recomendadas

Bovenberg, A. Lans, y Lawrence H. Goulder, "Environmental Taxation and Regulation", en *Handbook of Public Economics*, vol. 3, 1471 (Alan J. Auerbach y Martin Feldstein [comps.], 2ª ed., 2002).

Brickley, James A., Sanjob Misra y R. Lawrence Van Horn, "Contract Duration: Evidence from Franchising", 49 *J. Law & Econ.,* 173 (2006).

Bubb, Ryan, y Richard H. Pildes, *How Behavioral Economics Trims Its Sails and Why* (N.Y.U. School of Law, 25 de marzo, 2013).

Cost-Benefit Analysis: Legal, Economic, and Philosophical Perspectives (Matthew D. Adler & Eric A. Posner [comps.], 2001).

Crandall, Robert W., *Controlling Industrial Pollution* (1983).
Craswell, Richard, "Regulating Deceptive Advertising: The Role of Cost-Benefit Analysis", 64 *S. Cal. Rev.,* 549 (1991).
Crocker, Keith J., y Scott E. Masten, "Regulation and Administered Contracts Revisited: Lessons from Transaction-Cost Economics for Public Utility Regulation", 9 *J. Reg. Econ.*, 5 (1996).
Economics of Environmental Law (2 vols., Richard R. W. Brooks, Nathaniel O. Keohane y Douglas A. Kysar [comps.], 2009").
Handbook on the Politics of Regulation (David Levi-Faur [comp.], 2011).
Higgins, Richard S., y Fred S. McChesney, "Truth and Consequences: The Federal Trade Commission's Ad Substantiation Program", 6 *Int'l Rev. Law & Econ.,* 151 (1986).
Incentives for Environmental Protection (Thomas C. Schelling [comp.], 1983).
Laffont, Jean-Jacques, "The New Economics of Regulation Ten Years After", 62 *Econometrica,* 507 (1994).
Parry, Ian W. H., "Pollution Regulation and the Efficiency Gains from Technological Innovation", 14 *J. Reg. Econ.,* 229 (1998).
Peltzman, Sam, "An Evaluation of Consumer Protection Legislation: The 1962 Drug Amendments", 81 *J. Pol. Econ.,* 1049 (1973).
Posner, Richard A., *Catastrophe: Risk and Response,* cap. 3 (2004).
———, "The Federal Trade Commission: A Retrospective", 72 *Antitrust L. J.,* 761 (2005).
———, "Regulation (Agencies) *versus* Litigation (Courts): An Analytic Framework", en *Regulation* versus *Litigation: Perspectives from Law and Economics,* 11 (Daniel P. Kessler [comp.], 2011).
Regulation Versus Litigation: Perspectives from Law and Economics, 11 (Daniel P. Kessler [comp.], 2011).
Revesz, Richard L., "Environmental Regulation, Cost-Benefit Analysis, and the Discounting of Human Lives", 99 *Colum. L. Rev.,* 941 (1999).
Shavell, Steven, "A Model of the Optimal Use of Liability and Safety Regulation", 15 *RAND J. Econ.,* 271 (1984).
———, "Liability for Harm *versus* Regulation of Safety", 13 *J. Legal Stud.,* 357 (1984).
Sunstein, Cass R., *Risk and Reason: Safety, Law, and the Environment,* 154 (2002).
Viscusi, W. Kip, *Regulating Consumer Product Safety* (Am. Enterprise Institute, 1984).
———, y Joseph E. Aldy, "The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World", 27 *J. Risk & Uncertainty,* 5 (2003).
Zupan, Mark A., "The Efficacy of Franchise Bidding Schemes in the Case of Cable Television: Some Systematic Evidence", 32 *J. Law & Econ.,* 401 (1989).

Problemas

1. ¿Tendría el subsidio a una compañía que estuviera encaminado a sufragar el costo de reducir sus emisiones de contaminantes los mismos efectos de asignación y distributivos que una regulación que requiriera la reducción, sin compensación alguna? Sugerencia: ¿cuál sería el efecto del subsidio sobre la producción de la industria?

2. Si los costos de transacción entre los tenedores de *copyright* y las compañías de televisión por cable fueran bajos, ¿habrían tenido las reglas de la FCC sobre las señales distantes algún efecto sobre la producción de material con registro de marca? ¿Habría representado alguna diferencia que muchos de los *copyrights* hubieran sido propiedad de las cadenas de televisión, cada una de las cuales es dueña de varias estaciones televisivas?

3. Tal como se observó en el capítulo XI, los fabricantes de equipo de grabación doméstica de videos no enfrentan la exigencia de pagarles regalías de *copyright* a los propietarios de programas de televisión con registro de marca, una vez que se ha determinado que no han cometido una infracción contributiva. Los dueños de *copyright* han propuesto que el gobierno federal establezca un impuesto sobre la venta de grabadoras de video y casetes de video y que distribuya lo recaudado por el impuesto entre los propietarios de dichos derechos proporcionalmente a la parte que les corresponda de la audiencia televisiva, siendo las participaciones nuevamente calculadas de forma periódica. Evalúe las propiedades económicas del esquema propuesto en comparación con la responsabilidad del *copyright*.

4. Evalúe la siguiente proposición: Tenemos a la mano un excelente ejemplo de un impuesto regulatorio factible y efectivo sobre las externalidades dañinas para usarlo como un modelo para un impuesto a la contaminación: el impuesto específico a las bebidas alcohólicas.

5. Los fumadores acaso les adscribirían valores distintos a tres políticas, cada una de las cuales se espera que reduzca el número pronosticado de muertes debido a cáncer de pulmón por, digamos, la mitad: hacer a los cigarrillos más seguros para la salud, reducir la cantidad de fumadores y reducir el carácter letal del cáncer de pulmón. ¿Significa lo anterior que el análisis de costos-beneficios arroja una sola valuación de la trascendencia que tiene la vida de un individuo afecto al cigarro?

Cuarta Parte

EL DERECHO DE LAS ORGANIZACIONES MERCANTILES Y LOS MERCADOS FINANCIEROS

XV. CORPORACIONES, FINANCIAMIENTO ASEGURADO Y NO ASEGURADO, QUIEBRA[1]

XV.1. La naturaleza de la empresa

Los costos de transacción —los costos implicados en el ordenamiento de la actividad económica a través del intercambio voluntario— son un tema recurrente en este libro. Aquí lo emplearemos para explicar por qué una porción tan grande de la actividad económica es realizada por empresas y no por individuos.[2]

Se contrastan dos métodos de organización de la producción. En el primero, el empresario contrata con una persona la provisión de partes componentes, con otra el ensamblado de tales partes y con una tercera la venta del producto terminado. En el segundo, el empresario contrata a esas personas para que realicen estas tareas como sus empleados bajo su dirección. El primer método de organización de la producción es el dominio tradicional del mercado en economía y del derecho de los contratos en el derecho; el segundo es el dominio de la empresa en economía y del derecho de amo-sirviente en el derecho. La esencia del primer método es que el empresario negocia con cada uno de los tres productores un acuerdo que especifica el precio, la cantidad, la calidad, la fecha de entrega, los términos crediticios y las garantías del cumplimiento del contratista. La esencia del segundo método es que el empresario paga a los productores un salario: un precio no por un desempeño específico sino por el derecho a dirigir su desempeño. Los empleados que fungen de supervisores dirigen el trabajo de los empleados subordinados en lugar de comprar su producción como ocurre en un sistema de mercado.

Cada método tiene sus costos. El primero, el contrato, requiere que los detalles del desempeño del proveedor se expliciten en el momento de firmar el contrato. Esto podría requerir prolongadas negociaciones o elaborados procedimientos de regateo, y, si el cambio de las circunstancias requiriera la modificación de cualquier aspecto del desempeño convenido, el acuerdo deberá ser renegociado. El segundo método, la empresa, implica costos de incentivo, de información y de comunicación. Dado que al empleado (o equipo de empleados) no se le paga directamente por su producción, tiene menos

[1] Michael P. Dooley, *Fundamentals of Corporation Law* (1995); William A. Klein, J. Mark Ramseyer y Stephen M. Bainbridge, *Business Associations: Cases and Materials on Agency, Partnerships, and Corporations* (5ª ed., 2003).

[2] La explicación tiene sus raíces en otro conocido artículo de Ronald H. Coase, "The Nature of the Firm", 4 *Economica,* 386 (1937).

incentivo para minimizar sus costos. Y la información acerca de los costos y el valor está oculta en la empresa porque los empleados no regatean por los diversos recursos que utilizan en la producción, un proceso que indicaría dónde podrían emplearse más provechosamente tales recursos; en otras palabras, se pierde la información contenida en los precios. Además, dado que el desempeño de la empresa es dirigido por las órdenes del empleador, se necesita una maquinaria para minimizar las fallas de la comunicación en sentido descendente y ascendente a lo largo de la cadena de mando: una maquinaria que es a la vez costosa e imperfecta. En suma, el método del contrato para organizar la actividad económica se enfrenta con el problema de los costos de transacción elevados, el método de organizar la actividad económica por medio de la empresa se enfrenta con el problema de la pérdida de control.

Es el problema del control, o, como se le llama en la economía moderna, el problema de los costos de agencia (los costos para el principal de obtener de sus agentes un desempeño fiel y eficaz), antes que la ley de los rendimientos decrecientes, lo que limita el tamaño eficiente de las empresas.[3] Los rendimientos decrecientes sólo limitan la cantidad de un solo producto que puede producir una empresa eficientemente.[4] Debido a que hay un límite impuesto al lapso de supervisión a cargo de una persona, entre más empleados tiene la empresa, más supervisores necesita; y, entre más supervisores tiene, serán más los supervisores de los supervisores que requiera, ya que en cada nivel de la jerarquía se fija un límite al lapso en que ocurre el monitoreo. De este modo, conforme se expande la organización, los estratos de supervisores están destinados a multiplicarse. Los inevitables efectos de la jerarquía extendida que acompaña a una organización en expansión incluyen la demora en las órdenes ejecutivas, la pérdida de información, la recepción atenuada de las directrices procedentes de la cima y, en breve, el debilitamiento de la política de control y la merma de la coherencia. Y, entre más grande sea la organización, más difícil será establecer una correlación entre la labor de un empleado en particular y el valor de lo que se produzca en ella, por lo que los incentivos del empleado irán dejando de estar en consonancia con los de la empresa. A los empleados se les abrirán más espacios para adoptar un comportamiento apto para perseguir sus propios intereses pero no los de la compañía; en otras palabras, los costos de agencia se habrán de incrementar.

Los costos de agencia tienen muchas formas; la de bajar la guardia es solamente una. Otra es la de las "actividades dirigidas a influenciar": aquellas mediante las cuales los agentes tratan de influir en las decisiones de sus prin-

[3] Entre el contrato y la empresa se encuentra la relación de principal-agente —una relación fiduciaria, antes que distante—, cuando el agente no es un empleado del principal. Véase antes IV.8.

[4] Así, la "empresa" y la "integración vertical" (véase antes X.8) son sinónimos; ambos términos significan la sustitución de las negociaciones de mercado por la dirección jerárquica como el método para coordinar las distintas etapas de la producción-distribución.

cipales; por ejemplo, lisonjeándolos, o actuando como "el que a todo dice que sí" en vez de ser "el que a todo le pone peros", haciendo favores personales, estableciendo alianzas con los cotrabajadores y colocándose en una mejor posición. El agente puede incluso tratar de ayudar a su supervisor inmediato para derrocar a altos gerentes. Un costo de agencia relacionado con los previos es el de acaparar información a fin de volverse indispensable y entorpecer la producción de sus competidores dentro de la organización.

El grado en que se despliegan las actividades encaminadas a influenciar y las maniobras arribistas anexas puede guardar relación con el ritmo y la dirección del cambio registrado en el tamaño de una organización. Si ésta se está transformando, dichas actividades deben disminuir porque se brindan mejores oportunidades para la promoción. Si, por el contrario, la organización se está reduciendo, las actividades deben propagarse. Y ese ajuste a la baja propiciará esa búsqueda inescrupulosa de puestos elevados que va aparejada de un descenso en la productividad, con lo que menguará la capacidad de la compañía para competir en el mercado.

Dado que una empresa tiene el propósito de coordinar las actividades, seguramente su forma de organizarlas es más eficiente que la que consiste en dejar todo librado a las disposiciones y ordenamientos de carácter estrictamente privado —como los emitidos en los mercados, los acuerdos tácitos, las negociaciones, las concesiones y soluciones intermedias, las redes sociales, las costumbres, etc.—. Sin duda alguna, al parecer, un planificador central podría coordinar la actividad económica mejor de lo que lo puede hacer el mercado, con sus ciclos bursátiles, sus tasas de desempleo y demás indicios de desequilibrio. No obstante, como sostuvo acertadamente Friedrich Hayek,[5] y como lo confirmó el colapso del comunismo, esta idea enraizada en el sentido común es falaz. La razón fundamental es que el conocimiento, en especial el referente a la manera de hacer algo en concreto en lugar del que consiste en acumular en la memoria hechos o procedimientos formales que han de ser comunicados como un conjunto de instrucciones, es virtualmente intransferible. Como resulta de ello, es poco probable que el administrador de un sistema complejo tenga toda la información que requiere para tener la facultad de ejercer el control con inteligencia. De ahí la importancia de los métodos descentralizados de coordinación, tales como el del mercado económico, en donde el conocimiento disperso —el conocimiento de cada consumidor acerca de sus necesidades y oportunidades, el de cada vendedor acerca de sus costos, sus proveedores, sus clientes— viene a integrar un conglomerado merced al sistema de precios. Si, tras descubrir que existen probabilidades de que un cierto insumo requerido para la fabricación o la distribución de su producto escasee, un vendedor lo compra en grandes cantidades y lo almacena, su acción conduce por fuerza a que aumente el precio del in-

[5] F. A. Hayek, "The Use of Knowledge in Society", 35 *Am. Econ. Rev.*, 519 (1945).

sumo, y su precio más elevado induce a otros usuarios a economizar en su uso y, al incrementar de este modo sus propios costos, se ven impelidos a aumentar el precio que cobran por su producto, con lo que en efecto transmiten la información sobre los costos más altos del insumo al mercado entero. El precio opera como un método a través del cual la información privada es difundida a lo largo de un mercado y, en sentido último, a lo largo de la economía nacional y mundial en su totalidad. El precio acaba por hacer presa de la información.

Prescribir contractualmente lo que deben hacer los agentes en todos los estados potenciales del mundo tendría un costo prohibitivo. En vez de ello, las organizaciones se basan en un conjunto de normas comunes, formas de entenderse, costumbres y perspectivas que remplazan a la contratación explícita y, así, propician la cooperación respecto a los niveles de desempeño que no pueden ser prescritos por directrices formales. Este conjunto de elementos vinculantes ajenos a los lineamientos formales (en suma, la "cultura" de una organización) incluye códigos y otros tipos de capital humano específico compartido que facilitan la comunicación y la coordinación entre los agentes. Por desgracia, una cultura organizacional que es óptima en su ambiente actual puede trocarse en subóptima cuando el ambiente se modifica, y, aunque haya un esfuerzo de adaptación, ésta puede dificultarse porque, una vez que se han creado canales informativos y otros elementos organizacionales, la inversión realizada ha sido absorbida hasta tocar los cimientos y, por lo tanto, la respuesta de la organización a las nuevas circunstancias queda constreñida. El cambio resulta especialmente difícil de sobrellevar en virtud de que la cultura de una organización está diseminada a través de toda su estructura en lugar de estar concentrada en una sola área o foco (el constituido por un manual de trabajo, por ejemplo) donde pudiera ser cambiada de un plumazo. De ahí que exista una tendencia al conservadurismo organizacional.

XV.2. Las diferentes clases de empresas

La teoría de la empresa nos dice por qué se organiza tanta actividad económica en empresas, pero no por qué la mayoría de las empresas son corporaciones. Una clave es que las empresas en las que los insumos son primordialmente mano de obra antes que capital son a menudo asociaciones o propiedades individuales en lugar de corporaciones. La corporación es primordialmente un método para solucionar problemas que se encuentran al reunir cantidades sustanciales de capital.[6]

[6] Véase una taxonomía y un análisis útiles de las diferentes formas de la empresa en Henry Hansmann, "Ownership of the Firm", 4 *J. Law, Econ. & Organization,* 267 (1988); también Hansmann, *The Ownership of Enterprise* (1996). Las corporaciones sin fines de lucro y las cooperativas de trabajadores se analizan al final de esta sección.

¿Cómo reúne capital el empresario que tiene una idea prometedora para una nueva empresa, pero carece del capital necesario o no quiere arriesgarlo? Es probable que ni siquiera se considere tomar en préstamo todo el capital. Si la tasa de interés sin riesgo es 6%, pero la empresa tiene una probabilidad de 50% de fracasar y no tener activos para pagar el préstamo, el prestamista, si es neutral ante el riesgo, cobrará una tasa de interés de 112%.[7] La elevada carga de los intereses, más la amortización, impondrán a la empresa un costo fijo pesado desde el principio. Esto aumentará el peligro del fracaso (véase antes XIII.8) y, por ende, la tasa de interés.

Estas dificultades podrían superarse en principio mediante la elaboración cuidadosa e imaginativa del acuerdo de préstamo, pero los costos de transacción podrían ser muy elevados. Una alternativa es que el empresario admita a un socio en la empresa, el que tendría derecho a recibir una parte de sus ganancias, si las hay, a cambio de aportar a ella el capital necesario. La compensación del socio se determina automáticamente por la fortuna con que corra la empresa. No hay necesidad de calcular una tasa de interés, aunque esto está implícito al determinar la fracción de cualesquiera ganancias futuras que habrá de recibir el socio a cambio de su aportación. Es muy importante el hecho de que no haya costos fijos de deuda que vuelvan a la empresa más riesgosa de lo que es inherentemente; el socio recibirá ganancias sólo si se obtienen.

Todavía hay algunos problemas. Una asociación puede disolverse por cualquiera de los socios, y se disuelve automáticamente por la muerte de cualquiera de ellos. Es posible que la frágil permanencia del arreglo obstaculice el compromiso de grandes cantidades de dinero en una empresa en la que podría quedar congelado por años. Los socios podrían estar en posibilidad de negociar este obstáculo, pero no sin incurrir en costos de transacción que pueden ser elevados. Además, en la medida en que convengan en limitar el derecho del socio inversionista de disolver la asociación y retirar su dinero, la liquidez de su inversión se reduce y (como veremos más adelante en este capítulo) puede quedar a merced del socio activo. La inversión queda sellada, de un modo muy similar a como quien ha obtenido una franquicia puede quedarse atrapado en ella con una inversión comprometida, como la construcción de una instalación cuyo valor es mucho menor en cualquier uso alternativo.

Dado que cada uno de los socios es personalmente responsable de las deudas de la asociación, un inversionista potencial deseará saber la probable extensión de la responsabilidad potencial de la empresa, o aun participar en la administración efectiva de la empresa a fin de asegurarse de que no contraiga deudas mayores por las que él sería responsable. Y todavía subsistiría un riesgo de responsabilidad indefinida. En principio, la empresa podría in-

[7] Supongamos que el préstamo es de 100 dólares y que debe pagarse en un año. El prestamista deberá cobrar una tasa de interés que le genere un valor esperado de 106 dólares, dada una probabilidad de pago de 50%. Resolviendo la ecuación $0.5x = 106$ dólares para x nos da 212 dólares, es decir, 100 dólares de principal más 112 dólares de intereses.

cluir en todos sus contratos con clientes y proveedores una cláusula que limite su responsabilidad a los activos de la empresa (algunos fideicomisos hacen esto). Pero la negociación de tales dispensas sería costosa. Y sería en todo punto impracticable limitar de esta forma la mayor parte de la responsabilidad cuasidelictiva. Asegurarse tampoco es una respuesta completa, como veremos.

Los problemas de la asociación que hemos venido discutiendo surgen de la necesidad que tiene la empresa de un capital mayor que el que puede esperarse que aporten los trabajadores (el socio activo en nuestro ejemplo). Una empresa que no necesite capital "externo" se organizará probablemente como una asociación o como un solo propietario y no como una corporación, porque las ventajas de la forma corporativa se encuentran primordialmente en la superación de los problemas que encontraría una asociación o un propietario único al tratar de obtener capital de los inversionistas.

Otro tipo de empresa es la corporación sin fines de lucro: la forma común de las confesiones religiosas, colegios y universidades privados, y la mayoría de las otras instituciones de caridad. Dado que a una empresa sin fines de lucro se le prohíbe repartir cualquier exceso de ingresos sobre los costos como ganancias, una empresa que desee obtener ganancias no se organizará en una forma sin fines de lucro. Por lo tanto, el hecho de que una empresa haya escogido esa forma es una señal de que desea maximizar algo distinto de las ganancias, presumiblemente su producción. El punto no es que un precio privado de un componente de ganancia haya de ser inferior y por lo tanto suscite una mayor demanda del producto o servicio de la empresa. Probablemente no será en modo alguno más bajo, si es que la empresa debe compensar el capital invertido; pues no deja de ser un detalle secundario que el pago cobre la forma del interés, por un lado, o de los dividendos más la revalorización de los activos de la empresa, por otro. La clave para dar cuenta de la habilidad que posee una institución de beneficencia para vender su producto a precio módico (o para ofrecerlo gratuitamente) reside en que su capital es donado. El incremento resultante en la utilidad de los objetos procedentes de la caridad entra en las funciones de utilidad de los donantes; son altruistas, por lo que derivan un beneficio no pecuniario de sus donaciones que ocupa el lugar de un rendimiento financiero en su inversión y es el incentivo que subyace a la elección de la forma organizacional.

Así, los donantes de la corporación sin fines de lucro corresponden a los accionistas de una corporación lucrativa. Si los donantes no controlan a la junta directiva y ésta no necesita buscar nuevos donativos a fin de mantenerse en operación, se plantea el problema grave de los costos de agencia que se discuten antes, en XV.1. Pero si se requiere un flujo continuo de donativos y si, además, la corporación no lucrativa tiene consumidores que pagan —ambas condiciones que se satisfacen en los colegios y universidades—, los incentivos para la operación eficiente no parecen marcadamente menores que en el

caso de las corporaciones lucrativas, porque bajo estos supuestos la corporación sin fines de lucro enfrenta la competencia en el mercado de capital (donde es como una empresa de negocios que debe recaudar capital nuevo cada año, lo que la mayoría de las empresas de negocios no hacen) y en el mercado de productos.

Con todo, hay que tener cuidado en no destacar más de la cuenta el motivo altruista en las donaciones a las empresas sin fines de lucro. Las donaciones otorgadas a éstas obedecen a menudo al propósito de exhibir la riqueza del donante, atraerle publicidad, agenciarle la reputación de altruismo que lo hace merecedor de respeto o aun de admiración o bien promover fines políticos que, de ser alcanzados, lo beneficiarían en el ámbito financiero. Las donaciones hechas a colegios y universidades con frecuencia están motivadas por la esperanza de inducir a la institución a que admita a los hijos del donante. Los donantes pueden hacer aportaciones a esta o aquella institución sin fines de lucro dedicada a la investigación sobre enfermedades que sufren o temen contraer. Las donaciones hechas a las organizaciones religiosas pueden estar motivadas por el miedo a ser condenados o por la esperanza de alcanzar la salvación. Algunas donaciones a instituciones de caridad son trucos para evadir impuestos. Y los empleados de las empresas sin fines de lucro suelen tener una motivación idéntica a la de los que se desenvuelven en las lucrativas. Contemplada desde adentro, la administración de una universidad, por ejemplo, es una carrera en pos del éxito económico muy parecida a cualquier otra.

La asociación es un tipo de cooperativa de trabajadores, pero este término se usa más frecuentemente para describir a una empresa convencional intensiva en capital que es propiedad de los trabajadores. En una verdadera cooperativa de trabajadores, los trabajadores son los solicitantes residuales de cualesquiera ganancias de la empresa, pues el capital se reúne mediante una combinación de préstamos a tasas de interés fijas y de ganancias retenidas. La objeción tradicional es que la empresa estará siempre descapitalizada porque los trabajadores tienen un horizonte de tiempo más corto que el de la empresa. Un trabajador no desearía que se redujeran sus ingresos para financiar un proyecto que no generará ganancias antes de que él se jubile.

¿Podría haber algunas ventajas compensatorias en la forma de cooperativa? Se ha sostenido que "la protección al medio ambiente tiende a armonizarse mejor con los intereses y los ideales de las empresas administradas por los trabajadores que con los intereses de las empresas capitalistas", porque "los trabajadores, al revés de los capitalistas, tienen que vivir en las comunidades donde trabajan, de modo que deben vivir con la contaminación que generan".[8] Pero si "trabajadores" incluye a los oficinistas junto con los trabajadores fabriles, o si una empresa tiene varias o muchas fábricas y sólo algunas de

[8] Elizabeth Anderson, *Value in Ethics and Economics*, 213 (1993).

ellas contaminan, es posible que la mayoría de los trabajadores-propietarios no estuvieran expuestos a la contaminación de la empresa. Aun si están expuestos y tienen, por lo tanto, más qué perder de la contaminación que los accionistas de la empresa, tendrán también más que perder (sus empleos), por las medidas de control de la contaminación, que los accionistas. Al aumentar los costos de la empresa, tales medidas podrían volver a la empresa menos competitiva y obligarla a reducir su producción o aun a cerrar. Las cooperativas de triplay del noroeste, la principal "historia de éxito" de las empresas propiedad de los trabajadores en el sector industrial de los Estados Unidos, tienen, en opinión de un admirador de tales empresas, las mismas condiciones de trabajo sucias, ruidosas y peligrosas que los aserraderos capitalistas.[9] Además, prácticamente han desaparecido.[10]

El problema del horizonte acortado del trabajador sería menos grave si éstos conservaran sus acciones negociables o redimibles en su cooperativa. Ésta es la penetrante idea que subyace en el Plan de Propiedad de Acciones de los Empleados (ESOP), un instrumento que permite que el plan de beneficios de los empleados invierta sus activos en las acciones del empleador, el cual está organizado como una corporación lucrativa convencional. (El ESOP da empleo a unos ocho millones de trabajadores: muchos más que los reclutados en las cooperativas de trabajadores comunes y corrientes.) Merced a su fuerza, el valor de las acciones en las reservas del empleador que están en manos del ESOP se apodera de las futuras ganancias descontadas de éste, incluidas las ganancias habidas luego de que los empleados actuales renuncian o se retiran. De este modo, si tales ganancias se ven truncadas por la adopción de la clase de política de "empleo pleno" que sería atractiva para una cooperativa de trabajadores convencional, los trabajadores-propietarios resultan mermados en su capacidad como accionistas. Es cierto que el valor de su empleo puede exceder la disminución del valor de sus acciones que podría ser provocada por semejante política. Y, por consiguiente, la modalidad del ESOP diluiría, en lugar de eliminar, las ineficiencias propias de la cooperativa de trabajadores común y corriente… si no fuera porque ERISA (Ley de Seguridad de los Ingresos de Jubilación del Empleado) demanda que el ESOP sea administrado por un fiduciario que tenga prohibido considerar cualquier interés distinto del de los beneficiarios del plan en tanto que beneficiarios, es decir, como accionistas y no como empleados.[11]

El argumento a favor del ESOP (y el propósito declarado que siguen las exenciones de impuestos a la propiedad de acciones en el ESOP) es que, al dar a los trabajadores una participación más considerable en los caudales de la compañía, conduce a "ampliar la propiedad del capital dentro de un sistema

[9] Christopher Eaton Gunn, *Workers' Self-Management in the United States*, 130 (1984).

[10] Véase John Pencavel, *Worker Participation: Lessons from the Worker Co-Ops of the Pacific North West*, 20-25, 75 (2001).

[11] Véase Summers *vs.* State Street Bank & Trust Co., 104 F.3d, 105, 107-108 (7th Cir., 1997).

capitalista", cumpliendo así el objetivo de "fomentar la formación de capital a través de una mejora en el manejo de las finanzas, de una mejora en las relaciones administración-mano de obra y, en fin, de una mejor productividad".[12] ¿Qué es lo que la hace constituir "una mejora" en el manejo de las finanzas? Nada. Simplemente suministra un incentivo fiscal para la inversión y es, por ello, equivalente a reducir impuestos sobre las ganancias o retornos de capital. Pero puede suavizar las relaciones laborales al dar a los trabajadores una participación en las ganancias de la compañía, lo que llevaría a una moderación en sus demandas salariales; y podría operar como un instrumento para prevenir la absorción a manos de otra firma si los trabajadores se sintieran atemorizados ante la perspectiva de que ésta los despidiera o redujera ya sea los salarios o los beneficios.

La gran desventaja de la modalidad del ESOP reside en el grave riesgo de que incurra en un nivel inferior de diversificación, como veremos en el siguiente capítulo.

Algunas empresas que podrían organizarse como compañías de negocios convencionales pasan en cambio a tener como propietario al gobierno, que las echa a andar; antes, en XIV.12, consideramos el ejemplo de los caminos y carreteras de peaje y observamos por qué es probable que esas empresas sean menos eficientes que las compañías de negocios normales.

XV.3. La corporación como un contrato estándar

La forma corporativa es la solución normal que han desarrollado el derecho y la práctica mercantil para resolver los problemas de financiamiento y responsabilidad de las asociaciones. La existencia perpetua de la corporación elimina la necesidad de un acuerdo especial que limite el retiro o la disolución. La responsabilidad del accionista por las deudas de la corporación se limita al valor de sus acciones ("responsabilidad limitada"), y la inversión pasiva está protegida además por: *1)* un conjunto de derechos legales frente a los administradores y cualquier grupo de accionistas controlador, y *2)* el hecho de que la propiedad de una corporación queda denotada en los fragmentos de acciones de valor relativamente pequeño que pueden ser, y en el caso de las corporaciones más grandes son, negociadas en mercados organizados, lo que asegura (en el caso de las acciones negociables) la liquidez de los intereses de los propietarios. La forma corporativa permite que un inversionista realice pequeñas inversiones de capital a fin de reducir el riesgo mediante la diversificación (véase después XVI.1), y, si se trata de una compañía que cotiza en la bolsa, liquidar su inversión rápidamente y con poco costo. Sin la respon-

[12] Brett McDonnell, "ESOP's Failures: Fiduciary Duties When Managers of Employee-Owned Companies Vote to Entrench Themselves", 2000 *Colum. Bus. L. Rev.,* 199, 220-221 (2000).

sabilidad limitada no se permitiría siquiera que un accionista, como un socio, vendiera sus acciones sin el consentimiento de los demás accionistas, ya que si las vendiera a alguien más pobre que él aumentaría el riesgo de responsabilidad para los demás accionistas.

La responsabilidad limitada no es un instrumento para eliminar los riesgos del fracaso empresarial, sino para trasladarlos de los inversionistas a los acreedores voluntarios e involuntarios de la corporación: son ellos quienes soportan el riesgo del impago corporativo. Los acreedores deben ser pagados por soportar este riesgo.[13] ¿Por qué desearía y podría el inversionista desplazar una parte del riesgo de fracaso de la empresa al prestamista, dado que deberá compensarlo por soportar cualquier riesgo adicional? La respuesta es que el prestamista podría ser quien soporte mejor el riesgo. Por ejemplo, el banco está en una mejor posición que el accionista medio para hacer la evaluación del riesgo de un impago. El banco es un especialista en la evaluación de riesgos. El accionista podría saber poco o nada acerca del negocio en el que ha invertido y podría afrontar elevados costos de información para saberlo. Y es probable que el accionista, de ser individual, sienta mayor aversión hacia el riesgo que el banco, que puede eliminar o reducir grandemente el riesgo de la pérdida en un préstamo particular manteniendo una cartera de préstamos diversificada. Es cierto que los accionistas pueden contrarrestar cualesquiera riesgos en que incurra la corporación manteniendo una cartera de valores diversificada; pero esto sucede *porque* hay una responsabilidad limitada del accionista. Sin ella, incluso un accionista que mantenga una cartera diversificada de valores no estaría protegido contra el riesgo de verse obligado a entregar toda su riqueza para pagar las deudas de una de las corporaciones cuyas acciones posee. Se enfrentaría a un doloroso trueque entre tener una cartera subdiversificada y no poder (debido a la cantidad de acciones en la cartera) monitorear a las compañías por cuyas deudas podría ser responsable como accionista. Asimismo, el prestamista controla su exposición: su exposición máxima es el monto del préstamo. El accionista, si fuera responsable de las deudas de la corporación, no podría controlar su exposición porque no podría limitar el monto de los préstamos adeudados por la corporación.

[13] Para ilustrar, supongamos que, en nuestro ejemplo anterior de una empresa nueva con una probabilidad de éxito de 50%, los accionistas aportan la mitad del capital requerido y un banco aporta la otra mitad, y el banco estima que en caso de un fracaso serán suficientes los activos de la empresa para pagar 80% del principal del préstamo. Por lo tanto, su rembolso esperado es $0.5\ [K(1 + r)] + 0.5\ (0.8K)$, donde K es el principal del préstamo y r es la tasa de interés. Para que este pago esperado sea igual al principal más 6% de interés o, en otras palabras, para que $0.5\ [K(1 + r) + 0.8K]$ sea igual a 1.06K, r deberá ser de 32%, lo que es una tasa de interés muy elevada. Pero supongamos que si los accionistas comprometieran su riqueza personal en el pago del préstamo en el caso de que la empresa fracase, el banco consideraría el préstamo sin riesgo y cobraría sólo 6% de interés. La diferencia entre una tasa de interés de 6% y una de 32% es una medida de la compensación pagada al banco por los inversionistas de la nueva empresa por asumir una parte del riesgo de incumplimiento.

Concedamos que un accionista grande podría encontrarse en una mucho mejor posición para evaluar y monitorear el riesgo del impago de la corporación que un acreedor comercial pequeño. Sin embargo, este último puede protegerse de ordinario vendiendo a muchos clientes diferentes y limitando la extensión del crédito otorgado a un cliente en particular a un periodo corto. E incluso el gran accionista probablemente no monitorearía de un modo óptimo, porque sus esfuerzos estarían confiriendo un beneficio no compensado a los demás accionistas.

El argumento a favor de la responsabilidad limitada no es concluyente. Al disminuir los activos disponibles para pagar a los acreedores, la responsabilidad limitada aumenta los costos de los préstamos adquiridos por las corporaciones. Sin embargo, los accionistas pueden renunciar a su responsabilidad limitada. Por lo tanto, la responsabilidad limitada es sólo una regla de omisión, y es razonable si la mayoría de los accionistas no están dispuestos a renunciar a dicha responsabilidad a cambio de un mayor rendimiento esperado de su inversión (mayor porque los costos de los préstamos obtenidos por la empresa serían menores).

Podría parecer que la responsabilidad limitada permite a una empresa externalizar el riesgo del fracaso.[14] Sin embargo, el prestamista voluntario se ve plenamente compensado del riesgo del impago por la tasa de interés más elevada que la corporación deberá pagar a sus prestamistas en virtud de su responsabilidad limitada. También tiene la libertad de insistir, como una condición para hacer el préstamo, en que los accionistas garanticen personalmente las deudas de la corporación, o inserten en el acuerdo de préstamo otras disposiciones que limiten el riesgo del prestamista. Por supuesto, toda reducción resultante del riesgo del impago reducirá, no obstante, la tasa de interés; el acreedor que quiera tener más seguridad respecto al impago debe pagar por ello.

La tasa de interés reflejará el riesgo del impago tal como se estima ese riesgo cuando se firma el acuerdo de préstamos, y en consecuencia hay un peligro de que la corporación aumente más tarde el riesgo del impago, por ejemplo, obteniendo préstamos adicionales o transfiriendo activos a sus accionistas sin una consideración completa. Haciendo estas cosas, un prestatario reduce unilateralmente la tasa de interés que está pagando por el préstamo, una tasa negociada con referencia a un nivel de riesgo esperado que es menor que el que se produce efectivamente. Todo lo que aumente la razón de deuda a capital del prestatario aumentará la probabilidad del impago, porque las cargas de deudas son costos fijos para la empresa y no pueden reducirse si hay adversidades comerciales que reduzcan sus ingresos.

¿Por qué habría de aumentar la empresa deliberadamente el riesgo de impago? Este planteamiento de la cuestión es engañoso. La empresa *no quiere*

[14] Véase, por ejemplo, Jonathan M. Landers, "A Unified Approach to Parent, Subsidiary, and Affiliate Questions in Bankruptcy", 42 *U. Chi. L. Rev.,* 589, 619-620 (1975).

dejar de pagar. Quiere aumentar su rendimiento esperado; pero sólo puede hacerlo corriendo un riesgo de impago mayor. Supongamos que puede elegir entre dos proyectos de inversión que implican el mismo gasto. Uno tiene un rendimiento esperado de 25%, pero es riesgoso, y para financiarlo deberá obtener la empresa un préstamo a 15%, lo que bajará el rendimiento esperado neto de la empresa a 10%. El otro proyecto es menos riesgoso y tiene un rendimiento esperado de 18% (por lo que toca a la correlación positiva entre el riesgo y el rendimiento, véase el capítulo siguiente), y, por ser menos riesgoso, la empresa podría obtener un préstamo para el proyecto a 10%, de modo que el rendimiento neto esperado de ella es 2% menos (es decir, 8% contra 10% para el proyecto más riesgoso). Supongamos que 2% es todo el premio por el riesgo que demandan los propietarios de la empresa para invertir en el proyecto riesgoso, y supongamos ahora que la empresa obtiene prestado el dinero que requiere para cualquiera de los proyectos (recuérdese que cuestan lo mismo) de un prestamista que espera que el dinero se invierta en el proyecto 2. Al desviar el dinero de 10% al proyecto 1, el más riesgoso, la empresa puede aumentar a 15% su rendimiento neto esperado.

Esta estrategia no depende de la existencia de proyectos de inversión opcionales. Supongamos que una tasa de interés de 10% refleja un colchón de capital, esto es, resulta baja porque el riesgo del fracaso será soportado en gran medida por los accionistas. Eliminando el colchón tras obtener el préstamo, los prestatarios pueden aumentar su rendimiento esperado sin compensar al prestamista por el riesgo añadido. El rendimiento esperado será mayor porque el rendimiento se repartirá entre menos dólares del capital social. Supongamos que se espera que el proyecto genere una ganancia de un millón de dólares. Si hay 10 millones de capital social, esto se traduce en un rendimiento de 10%; si sólo hay 5 millones (porque los accionistas se han devuelto a ellos mismos la mitad en dividendos y han pedido prestado para reponerlo), esto se traduce en un rendimiento de 20 por ciento.

Esta maniobra fracasaría si los prestamistas originales pudieran elevar su tasa de interés para reflejar el riesgo añadido (véase después XVI.2). Pero, dado que la probabilidad de que el prestatario intente deliberadamente aumentar el riesgo del préstamo sería difícil de cuantificar, es poco factible que el prestamista se base para su protección en una tasa de interés más alta… en especial si la tasa más elevada incrementara el riesgo de impago, al generar un incremento en los costos fijos del prestatario. Para ganar acceso al mercado de préstamos a tasas razonables bajo tales circunstancias, podría ser del interés del prestatario corporativo disociarse de tal comportamiento oportunista llegando a un acuerdo que restrinja su libertad de acción; por ejemplo limitando su endeudamiento total y el monto de los dividendos pagaderos durante el tiempo de aplicación de las condiciones del préstamo, donde "dividendo" se define ampliamente para estos propósitos a fin de incluir cualquier transferencia de los activos corporativos por menos que su valor de mercado

completo. O bien el prestamista podría convenir en mantener alguna capitalización mínima especificada.

Como vimos en el capítulo IV, una de las funciones del derecho de los contratos es la de economizar en los costos de transacción proporcionando términos contractuales estándares que las partes tendrían que adoptar de otro modo mediante un acuerdo expreso, y ya debe estar claro que lo mismo se aplica al derecho de las corporaciones. Por ejemplo, las reglas del derecho de las corporaciones que limitan el pago de dividendos a la cantidad de excedente ganado que aparezca en los libros de la corporación otorgan a los acreedores cierta protección contra los intentos de los prestatarios corporativos de aumentar el riesgo de impago después de realizado el préstamo: una protección que, como hemos visto, en ausencia de una regla legal los acreedores escribirían probablemente en forma expresa en cada contrato de préstamo.

La analogía con el contrato desaparece en el caso de las extensiones de crédito *involuntarias*, como ocurre cuando un peatón es golpeado por un camión de mudanzas en movimiento en circunstancias que hacen a la compañía de mudanzas responsable de un cuasidelito ante el peatón. Dado que las partes no tuvieron ninguna oportunidad de transigir por adelantado acerca de las provisiones del derecho de las corporaciones, el peatón no será compensado por soportar el riesgo de impago creado por la responsabilidad limitada de la compañía de mudanzas.

Aun cuando el contexto sea el de una transacción voluntaria, los costos de la negociación explícita del alcance de la responsabilidad podrían ser elevados en relación con los daños implicados. La escasa probabilidad de que un empleado se vea gravemente lesionado en el trabajo, cuando se multiplica por la probabilidad adicional de que el empleador carezca de activos suficientes para satisfacer la reclamación del empleado por la compensación de los trabajadores, podría ser demasiado pequeña para ameritar la inclusión de una cláusula expresa en el contrato de empleo para cubrir esa contingencia (por ejemplo, una fianza del empleador). Cualquier término que sea implicado como una cuestión del derecho corporativo o del derecho de quiebras controlará las relaciones de las partes, aunque sea contrario a lo que habrían negociado las partes en un mundo de costos de transacción nulos. Pero hay una diferencia entre esto y el caso anterior del accidente: la tasa salarial puede ajustarse para compensar al trabajador por el riesgo de impago de cualquier derecho de compensación que pueda tener algún día contra su empleador. No habrá compensación para el peatón, porque no tiene una relación contractual real o potencial con la corporación.

Dado que la responsabilidad limitada podría generar costos externos porque una corporación podría tener acreedores involuntarios (no sólo las víctimas de cuasidelitos sino también las autoridades fiscales y las reguladoras), y que la quiebra es una fuente de costos sociales y no sólo un método para la transferencia de riqueza a los acreedores y entre ellos, podría pensar-

se que el gobierno debiera asumir un papel más activo para asegurar la solvencia de las corporaciones. Esto podría lograrse requiriendo que una corporación mantenga una razón fija de capital social a pasivos y limitando el derecho de la corporación a emprender proyectos riesgosos. Éste es el régimen establecido en el derecho bancario (discutido en el capítulo siguiente) y en el derecho corporativo europeo. Sin embargo, en ambos casos hay un escrutinio regulador continuo de la corporación a manos de una dependencia administrativa (aunque ésta no logró impedir el colapso financiero de 2008, del que se ocupa el próximo capítulo), una solución del Estado que hasta ahora ha sido resistida en la mayoría de las industrias no financieras de los Estados Unidos.[15]

XV.4. DEUDA CORPORATIVA Y DEUDA CON RESPALDO ESPECÍFICO

Hasta ahora hemos dado por sentado que las corporaciones tendrán deuda al igual que capital social: acreedores y accionistas. Hasta cierto punto, esto es realmente inevitable; una corporación no puede evitar la posibilidad de convertirse en un acreedor involuntario a resultas de un cuasidelito cometido por un empleado en el curso del desempeño de su trabajo. Pero ése es el tipo de riesgo contra el cual se puede proteger fácilmente mediante un seguro (aunque no por completo, como veremos). Si la deuda corporativa no es esencial, ya no resulta claro por qué es necesaria la responsabilidad limitada. Porque si las corporaciones tuvieran sólo capital social, de modo que no pudieran dejar de pagar sus préstamos, lo único en riesgo, a menos que la corporación cometiera un ilícito, serían las aportaciones de capital de los accionistas; y una responsabilidad cuasidelictiva no es asegurable.

Hay varias explicaciones de que las corporaciones tengan deuda al igual que capital social en sus estructuras de capital, aunque una de tales razones ha sido cuestionada: a saber, la de que así se incrementa la tasa de rendimiento de los accionistas al volver apalancada su inversión. Examinamos este cuestionamiento (y su refutación) en el siguiente capítulo.

Otra explicación es que una mezcla de capital accionario y capital de deuda permite que la corporación provea paquetes de riesgo y rendimiento diferentes para satisfacer las diversas preferencias de los inversionistas (un término empleado aquí en sentido amplio para incluir a los prestamistas y a los accionistas). Una corporación que tiene sólo capital accionario proveerá una inversión que tiene un riesgo y un rendimiento esperados uniformes para todos los inversionistas. No obstante, en cuanto la corporación toma prestado algún dinero, aumentarán el riesgo y el rendimiento esperados por los accionistas,

[15] Véase Geoffrey P. Miller, "Das Kapital: Solvency Regulation of the American Business Enterprise", en *Chicago Lectures in Law and Economics,* 61 (Eric A. Posner [comp.], 2002).

quienes tendrán ahora una inversión apalancada. Pero, aun si la razón de deuda a capital social es muy elevada, los acreedores tendrán una inversión menos riesgosa que la de los accionistas en una compañía que sólo tiene capital social por acciones, porque los acreedores no pierden nada de su inversión hasta que la inversión de los accionistas se pierda por completo. Los acreedores con pagos respaldados en activos se encuentran en una posición mejor aún: sólo tienen que preocuparse por el valor de los activos específicos destinados al pago de sus préstamos; a menos que ese valor baje más allá del valor de sus préstamos, esos acreedores no tienen por qué preocuparse aunque la corporación se vaya a la quiebra. ("Los bienes retenidos específicamente como garantía de pago no se ven afectados por la quiebra", como dicen los abogados… con cierta exageración, empero, puesto que hacer cumplir un derecho de retención no está exento de costos; y, si el valor del bien retenido es menor a la reclamación del acreedor, éste querrá levantar una reclamación por bancarrota con la esperanza de que una parte excedente de la reclamación sobre el valor del bien retenido se podrá recuperar con base en los activos no asegurados del deudor.) Esto quiere decir que los acreedores con garantía de pago se ahorran los costos del monitoreo del desempeño de la corporación; no tienen que evaluar a la empresa; sólo tienen que evaluar la garantía. Sin embargo, su rendimiento esperado es en consecuencia menor que el de los acreedores no asegurados mediante garantías de pago, quienes soportan costos de riesgo y de monitoreo mayores, y es mucho menor que el de los accionistas.

La presencia de otros acreedores vuelve desconcertante el fenómeno del préstamo con garantía específica en el contexto corporativo. La seguridad reduce los costos de evaluación del riesgo y del monitoreo para un acreedor, así como reduce la tasa de interés que él demande. Pero al mismo tiempo, al reducir el conjunto de los activos disponibles para satisfacer las reclamaciones de otros acreedores, la asignación de los activos al acreedor garantizado eleva la tasa de interés que el prestatario debe pagar a los no garantizados. Es importante la frase de "en el contexto corporativo". El préstamo con pago respaldado no es un enigma en el contexto del consumidor, por ejemplo, en el financiamiento de una casa o de un automóvil. Es improbable que el consumidor tenga otros acreedores grandes. No ocurre lo mismo con la corporación. Entonces, ¿por qué se ve ésta en mejor situación si pide prestado con la garantía de bienes específicos?

La respuesta podría encontrarse en la diferencia existente entre los costos de evaluación del riesgo y los de monitoreo (costos de información del acreedor), por una parte, y la pérdida esperada del acreedor en caso de impago, por la otra parte. El crédito asegurado con la garantía reduce los costos de información de los acreedores con seguridad de pago sin aumentar, y posiblemente aun reduciendo, los costos de los acreedores que no tienen esa seguridad. Si todos los créditos carecen de garantía, ningún acreedor sabe

dónde se encuentra. El primer prestamista que otorga un crédito a la corporación ignora el riesgo de impago porque no sabe cuánto dinero pedirá prestado la corporación a otros prestamistas. Cada acreedor subsiguiente se encuentra en la misma posición menos el último (sabrá que es el último si el préstamo limita la capacidad del prestatario de obtener crédito adicional, lo que minaría la solvencia del prestatario). Pero supongamos que algunos acreedores reciben en garantía intereses de los activos de la corporación, los que pueden embargar y vender en caso de impago, pagándose con el dinero de las ventas. Estos acreedores tendrán una idea mucho mejor del riesgo de no poder cobrar el dinero que se les debe. Sus costos de información serán bajos, mientras que los costos de información de los acreedores que no tienen garantía no serán mayores que antes, y quizá sean menores, porque el conjunto de los acreedores sin garantía es menor y el conjunto de los activos disponibles para satisfacer a los acreedores no asegurados es también menor.

El punto es que los costos de información de seguridad y de los acreedores son sustitutos. Si los costos de información varían entre acreedores, algunos querrán seguridad y otros, quizás especialistas en evaluación y monitoreo de riesgo,[16] estarán conformes con prescindir de ella.

Existe el peligro de que una corporación asuma demasiada deuda garantizada. La razón es que algunos acreedores no asegurados no pueden protegerse contra el mayor riesgo de impago creado para ellos por el crédito asegurado. Los acreedores involuntarios son el ejemplo obvio, pero los acreedores voluntarios con elevados costos de información crediticia constituyen otro. Utilizando las cifras del párrafo precedente, podemos ver que, si la corporación que tiene inicialmente sólo acreedores no asegurados pudiera sustituir con deuda garantizada la mitad de su deuda no garantizada sin tener que pagar una prima mayor por el riesgo de impago a sus demás acreedores no garantizados, estaría en mejor situación aunque no hubiese ninguna reducción neta de los costos de información de los acreedores… es más, aun si tales costos

[16] Por ejemplo, los acreedores comerciales podrían ser expertos en la evaluación de la salud financiera general del deudor, mientras que el vendedor de equipo de capital a crédito podría ser muy superior en la evaluación de la probabilidad de la recuperación de su préstamo si se ve obligado a demandar por el impago del deudor. Simplemente, los acreedores comerciales podrían tener tan escasa deuda vigente de cualquiera de sus clientes que el riesgo de impago sea un costo menor de sus operaciones. Debe añadirse que el crédito garantizado vuelve más fácil que una empresa inicie nuevos proyectos; las perspectivas de un nuevo proyecto son inciertas de manera inherente, así que tener asegurada la forma de ahorrar costos de información del acreedor es lo óptimo cuando se trata de un préstamo para lo que se tiene en mente. La explicación que se ofrece en el texto para el crédito garantizado ayuda a explicar la razón de que las empresas otorguen a menudo intereses de garantías en blanco sobre todos sus activos y de que las estructuras del capital de las corporaciones se hayan vuelto mucho más simples a lo largo del tiempo. Tanto las garantías en blanco como las estructuras corporativas simples reducen los costos de información de los acreedores. Si todos los activos están garantizados, los acreedores no garantizados no tienen por qué preocuparse acerca del valor de cualquiera de los activos: ninguno de éstos estará disponible para rembolsar los préstamos de los acreedores en el caso de un impago.

aumentaran (siempre que el aumento fuese menor de la mitad de 1%). Se ha sugerido que, a fin de impedir que las corporaciones asuman cantidades ineficientes de deuda garantizada, debiera modificarse la prioridad absoluta de los acreedores asegurados sobre los acreedores no asegurados.[17]

La deuda corporativa reduce los costos de evaluación del riesgo y del monitoreo en otra forma. El rendimiento del prestamista se fija por adelantado, de modo que sólo tiene que preocuparse por la insolvencia de la corporación; no tiene que preocuparse por saber si la corporación logrará maximizar sus ganancias. La deuda corporativa permite así la inversión en el sector corporativo por inversionistas a quienes les resultaría demasiado costoso monitorear a profundidad el desempeño empresarial. Sin embargo, esto se aplica principalmente a las corporaciones pequeñas, porque en las corporaciones grandes los accionistas también son, como sabemos, inversionistas pasivos, protegidos como están en su pasividad por la responsabilidad limitada y por la capacidad para diversificar sus propiedades entre diversas corporaciones. Pero esta última protección no impide que los accionistas sean "atrapados" por los administradores corporativos, un problema del que nos ocuparemos pronto.

Podemos invertir la interrogante y preguntar por qué una corporación *tiene* accionistas en vez de sólo acreedores, puesto que los inversionistas son sólo un tipo de acreedor. La respuesta es que hay una incertidumbre inevitable acerca del nivel de las ganancias, y los accionistas son los inversionistas que han aceptado ser los receptores de esta corriente incierta de ganancias en lugar de negociar un rendimiento fijo. Aceptan un riesgo más alto a cambio de un rendimiento esperado más elevado, y pueden aminorar el riesgo manejando una cartera diversificada de valores. Al contar con accionistas y acreedores (así como con diferentes tipos de acreedores), la corporación satisface las diferentes preferencias de personas que tienen dinero para invertir (o para prestar, que es una forma de inversión).

El estatus de los accionistas como el de quienes soportan el riesgo residual debiera ayudarnos a entender la razón de que los administradores efectivos de la corporación —la junta directiva y los ejecutivos que contrata y supervisa— sean escogidos (por lo menos mientras no surja la insolvencia) por los accionistas y no por los acreedores. Los acreedores no tienen el incentivo económico correcto —para maximizar el valor de la corporación— porque no tienen ningún derecho al extremo superior de la distribución de las ganancias posibles. Supongamos que a los acreedores se les deben 100 dólares y que hay una estrategia corporativa que hará valer a la corporación 101 con certeza, pero también otras que la harán valer 90 con una probabilidad de 10% y 200 con una de 90%. El valor esperado de la corporación es mucho mayor en el segundo escenario —189 ((90 × 0.1) + (200 × 0.9) = 189)

[17] Lucian Arye Bebchuk y Jesse M. Fried, "The Uneasy Case for the Priority of Secured Claims in Bankruptcy", 105 *Yale L. J.,* 857 (1996).

contra 101—, pero, si los acreedores estuvieran administrando la corporación, optarían por la primera estrategia.

Éste es un problema de los costos de agencia, y se reduce pero no se resuelve completamente si se otorga el control de la corporación a los accionistas. Es posible que ningún accionista tenga un incentivo para gastar tiempo y dinero en el monitoreo del desempeño de quienes administran la compañía, porque el beneficio de sus esfuerzos iría en su mayor parte a las manos de los otros accionistas. Esto sería importante a pesar de que fuera pequeño el costo del monitoreo, porque, aunque en ese caso podría el beneficio de un solo accionista superar ese costo, cada uno de los accionistas se vería tentado a abstenerse con la esperanza de que otro accionista se hiciera cargo del asunto. El problema del gorrón se agrava si los costos del monitoreo superan los beneficios de cualquiera de los accionistas, pero son menores a los beneficios para todos ellos. Y es probable que esto ocurra, porque los costos de monitorear a los administradores de una corporación grande son sustanciales.

Un procedimiento para controlar a los administradores sin que los monitoreen los accionistas consiste en contratar personas que vigilen a los administradores. Las juntas directivas y los auditores se encargan de esto; también lo hacen los grandes tenedores de deuda corporativa, porque al prestar dinero a la corporación obtienen un interés en su éxito. Aunque éste es un interés menor que el de los accionistas, podría concentrarse en menos personas, con lo que el problema del gorrón se vuelve menos agudo.

Otro procedimiento para controlar a los administradores consiste en mantener un riesgo importante de quiebra, lo que usualmente requiere la existencia de la deuda porque sin ésta los costos fijos de una corporación tienden a ser bajos, facilitando que la empresa se ajuste a una caída repentina de la demanda al reducir sus insumos (es decir, sus costos variables). La quiebra puede imponer costos a los administradores en una de dos formas. Primero, la quiebra es un símbolo muy visible del fracaso de una empresa, y, aunque no todas las quiebras se deben a los errores, la incompetencia o la deshonestidad de los administradores —algunas de ellas son la consecuencia inevitable de asumir racionalmente el riesgo—, los empleadores potenciales del administrador anterior de una empresa quebrada podrían considerar costoso determinar su culpa personal, de modo que descartarán sus esfuerzos por exculparse. Esto es más probable entre mayor sea la fracción de las quiebras que se deben, *efectivamente*, en todo o en parte, al error gerencial evitable.

El problema de la culpa excesiva se ve agravado por el "sesgo retrospectivo", el obstinado defecto cognitivo que lleva a las personas a pensar que algo era inevitable, como un fracaso empresarial, que *ex ante* podría haber tenido una probabilidad muy baja. Dado el sesgo retrospectivo, debemos considerar si el riesgo de la quiebra pudiera hacer que los administradores se volvieran *demasiado* cautelosos, en vista de que se les podría culpar por haber asumido sensatamente riesgos que, empero, acabaron por producir un desenlace desa-

fortunado (eventualidad que es intrínseca a la naturaleza de todo riesgo). Así pues, es probable que la quiebra corporativa resulte en particular costosa para un administrador si tiene un capital humano específico de la empresa y la quiebra conduce ya sea a la liquidación de la firma o al despido del administrador o a un castigo de otra índole.

El endeudamiento es una fuente de capital más barata que el capital accionario a causa del impuesto al ingreso corporativo, que es en parte un impuesto indirecto al capital accionario (véase después XVIII.5).[18] Ésta no es una buena razón económica para otorgar a la deuda corporativa un tratamiento fiscal favorable. La discriminación contra el capital accionario simplemente aumenta el riesgo de quiebra. La quiebra se ve precipitada por la deuda; una empresa que nunca obtuviera préstamos no podría irse a la quiebra a menos que se viera afectada por la deuda involuntaria, como en el caso de una condena de cuasidelito. Otro efecto de la discriminación contra el capital accionario es el debilitamiento del control de los accionistas sobre la corporación al alentar la retención de ganancias, lo que aumenta la discreción gerencial, antes que la repartición de ganancias entre los accionistas como dividendos. Se alienta la retención porque las ganancias que se reparten como dividendos están gravadas dos veces: como ingreso corporativo cuando se obtienen, y como ingreso individual cuando se reparten.

XV.5. LA QUIEBRA CORPORATIVA

A primera vista, la "quiebra corporativa" parece un oxímoron. Dado que los accionistas no pueden verse arruinados por el colapso de la corporación —su pérdida está limitada a lo que hayan invertido—, parecería que no habría necesidad de una quiebra corporativa ya sea voluntaria o involuntaria. Pero ésta es una concepción demasiado estrecha de los propósitos de la quiebra. Es cierto que una función importante de la quiebra *personal* es la de alentar la actividad empresarial de individuos que sienten aversión por el riesgo,[19] mediante el principio de que la quiebra libera a la persona en bancarrota de sus deudas, o por lo menos de la mayor parte de ellas. De otro modo, un individuo tendría que arriesgar toda su capacidad para obtener ingresos en sus proyectos comerciales. La responsabilidad limitada es a los empresarios corporativos lo que el derecho a declarar la quiebra personal es a los empresarios individuales.

Pero la quiebra es un remedio para el acreedor tanto como un derecho del deudor. Si un deudor carece de efectivo para pagar una deuda cuando se

[18] Véase más adelante, en XVI.2, una discusión más a fondo de la estructura del capital de la corporación.

[19] John Armour, "Bankruptcy Law and Entrepreneurship", 10 *Am. Law & Econ. Rev.,* 303 (2008).

vence, y un acreedor insiste en que se le pague, tiene que haber un mecanismo para el cobro o la cancelación aunque el deudor no tenga otros acreedores. En el caso de un deudor individual por oposición a uno corporativo, la interrogante crítica consiste en saber cuáles activos del deudor puede tomar el acreedor. Entre más activos pueda tomar el acreedor, más riesgosa será la actividad empresarial para el empresario; entre menos sean los activos que pueda tomar el acreedor, mayores serán las tasas de interés. Algunos estados de la Unión Americana tienen generosas exenciones hogareñas para los deudores insolventes, mientras que otros tienen sólo exenciones miserables.[20] En los primeros estados, el riesgo de ser empresario se reduce porque el costo del fracaso es menor, pero las tasas de interés son mayores porque el impago es más probable y la posición del acreedor en el caso del impago es más débil. A su vez, las tasas de interés más elevadas vuelven más probable el impago. Pero por otra parte tenemos el hecho de que, en los estados con exenciones bajas, es menor el riesgo de los prestamistas, lo que los induce a hacer préstamos más riesgosos; es decir, préstamos que más probablemente conducirán a la quiebra. Por lo tanto, no está claro, como una proposición teórica, si habrá más quiebras en los estados con exenciones altas que en los que tienen exenciones bajas.[21]

Cuando el deudor tiene varios acreedores, la quiebra se convierte en un remedio para el acreedor de igual importancia en el contexto corporativo que en el personal. Supongamos que el mercado del único producto de la corporación se derrumba, y ahora el precio de mercado es menor a los costos variables de la corporación a cualquier nivel de producción, de modo que lo único sensato que queda por hacer es cerrar y, suponiendo que no hay ninguna perspectiva de recuperación, vender los activos de la corporación a su valor residual o valor de salvamento. Si ese valor es menor a la deuda total de la corporación, los accionistas serán indiferentes a las oportunidades generadas por la liquidación. Su decisión racional será el abandono de la compañía. Los acreedores tendrán un interés en la liquidación de los activos de la corporación al mayor valor posible, pero se encontrarán con externalidades al tratar de elaborar un plan de liquidación que logre esta meta. ¿Qué diremos si el pago de la renta se adeuda y, si no se paga, los activos de la compañía tendrán que ser mudados a un costo grande? ¿Cuál de los acreedores pagará por ello? Cada uno tendrá un incentivo para hacerse el remolón, con la esperanza de que otro se adelante. Peor aún, aunque no haya dinero suficiente para

[20] La Ley de Quiebras es federal, pero permite que el deudor utilice las exenciones de su estado en lugar de las exenciones federales uniformes provistas en el Código de Quiebras.

[21] Un estudio encuentra que por cada 10 000 dólares de incremento en las exenciones hacendarias de un estado se registra en las quiebras de los negocios un aumento de 8%, y que por cada incremento de mil dólares en la excepción a la propiedad personal en un estado éstas aumentan 4%. Sumit Agarwal *et al.*, "Impact of State Exemption Laws on Small Business Bankruptcy Decision", 71 *S. Econ. J.,* 620 (2005).

el pago de la renta en la cuenta bancaria de la compañía, uno de los acreedores podría embargar la cuenta y tomar el dinero para cobrar su deuda, dejando que los demás acreedores se preocupen por la renta. La Ley de Quiebras se ocupa de este problema otorgando la primera prioridad a quien preste al quebrado el dinero indispensable para la preservación del valor de sus activos.

Otra posibilidad en un mundo sin la Ley de Quiebras consiste en que algunos acreedores hagan tratos personales con los accionistas que perjudicarían a otros acreedores. Veamos un ejemplo extremo (una cesión fraudulenta). Supongamos que la compañía tiene deudas por un total de un millón de dólares, del cual se adeudan 100 000 al acreedor A, y el valor total de los activos de la compañía es de sólo 200 000. El acreedor A podría ofrecer a la compañía comprarle los activos en 100 000 dólares, ya que podría revenderlos por 200 000 y recuperar su préstamo en esa forma. Los accionistas se sentirían felices porque estarían recuperando una parte de su inversión, pero los demás acreedores se quedarían sin nada. La posibilidad de tales tratos induciría costosas trampas entre los acreedores; eso y la incertidumbre de los rendimientos de los acreedores en tal régimen generarían mayores tasas de interés para compensar el riesgo mayor que los acreedores estarían corriendo.

Existiría este tipo de problema aun cuando la empresa simplemente le pagara al acreedor A su deuda completa antes que a los demás acreedores (quizá porque A fue el primero en reducir su pretensión de acudir a los tribunales); es decir, si le otorgara preferencia a A en el idioma de la Ley de Quiebras. El efecto sobre los otros acreedores sería la reducción de sus perspectivas de pago, de 20 centavos por dólar (200 000 divididos entre un millón) a 11 centavos por dólar (100 000 dólares divididos entre 900 000 dólares), y la intensificación de los esfuerzos de cada uno de esos acreedores para obtener una preferencia. En un mundo así, cada acreedor tendrá un incentivo para ser el primero en obtener un fallo en contra de la persona en bancarrota, y es probable que la carrera prive a la compañía de sus activos con demasiada rapidez para maximizar el valor de tales activos en una venta preparada con todo orden. (Sería como una estampida bancaria.) Cada acreedor tendría un incentivo para moverse mucho muy rápido —desde el punto de vista tocante a la maximización del valor de los activos del deudor— al buscar la satisfacción de su reclamo, y mucho muy despacio al pagar cualesquiera gastos requeridos para maximizar ese valor.

Estos problemas podrían resolverse mediante contratos *ex ante* entre los acreedores y entre éstos y los deudores. La Ley de Quiebras puede verse como el estándar de ese contrato. Los acreedores y los deudores no tienen que negociar acerca de lo que usualmente es, después de todo, un evento de probabilidad baja. En los términos de la Ley de Quiebras, una persona neutral designada por el tribunal, el fiduciario de la quiebra, desecha toda preferencia de último minuto y administra los activos de la persona en bancarrota como el representante de todos los acreedores (no asegurados). Como un administrador fiduciario convencional en un fideicomiso con varios beneficiarios, el fi-

duciario de la quiebra supera la dificultad que se crea cuando varias partes reclaman intereses en los mismos activos. Analíticamente, el problema de la quiebra, considerado desde el punto de vista de los acreedores, es el mismo que el de la propiedad dividida de la tierra (véase antes III.10).

La forma más interesante de la quiebra *corporativa* voluntaria es la reorganización corporativa. En los términos del capítulo 11 del Código de Quiebras, se permite que los administradores continúen operando la corporación como deudores en posesión; no hay administrador fiduciario, y no se dan pasos para liquidar la empresa. Pero la corporación debe elaborar en el término de seis meses un plan de reorganización, de modo que la empresa continúe operando, pero con una estructura de propiedad diferente. El plan típico propondrá que se convierta la deuda de la corporación en acciones y otros valores que se asignarán a los acreedores en pago de la deuda que la corporación tiene con ellos, así que se convertirán en propietarios de la corporación. Los acreedores pueden objetar los términos de la propuesta, y entonces corresponderá al tribunal de quiebras determinar el valor verdadero de la corporación y la forma como deberá repartirse, en valores de la compañía reorganizada, entre los acreedores.

Podría parecer que, si la continuación de la operación de la empresa fuese económicamente viable, los acreedores no tendrían ningún interés en forzarla a la quiebra en primer lugar.[22] Pero esto no es cierto. Una empresa puede ser a la vez insolvente y económicamente viable. Si la demanda del producto (o los productos) de la empresa ha declinado inesperadamente, la empresa podría descubrir que sus ingresos no cubren sus costos totales, incluidos los costos fijos de la deuda. Pero pueden exceder a sus costos variables, en cuyo caso no debiera ser liquidada, al menos no todavía, y es posible que a largo plazo pudiera la empresa continuar en su negocio indefinidamente con una planta más pequeña, pues en ese caso quizá no tendría que remplazar toda su deuda cuando esa deuda se retirara, sus costos totales serían menores, y sus curvas de demanda (menor) y de oferta (menor) podrían intersecarse de nuevo. En suma, es posible que la compañía tenga un futuro viable, a corto o largo plazo, al que podrá llegar si puede liberarse de su deuda actual. Un procedimiento

[22] Aunque la administración asume la iniciativa en la proposición del plan de reorganización, raras veces demanda la protección en los términos del capítulo 11 (la provisión de la reorganización), a menos que los acreedores estén a punto de forzar la quiebra de la compañía. La excepción principal ocurre cuando la compañía está afrontando una enorme responsabilidad cuasidelictiva *futura*. Quizás acaba de descubrir que su producto ha perjudicado a un millón de personas, pero éstas no han demandado aún, de modo que la compañía no tiene dificultad para pagar sus cuentas y quizá no la tenga durante muchos años. La compañía podría desear la reorganización a fin de proveer ahora para sus responsabilidades futuras (al considerar como acreedores a los futuros reclamantes potenciales por un cuasidelito y darles en la empresa reorganizada valores a cambio de anular sus reclamos, tal como si se tratara de acreedores comunes), de modo que pueda conducir sus actividades sin la incertidumbre creada por pasivos futuros grandes no pagados.

para lograrlo consiste en convertir esa deuda en capital accionario, en cuyo momento la deuda dejará de ser un costo fijo y dejará de impedir que la compañía sufrague sus otros gastos. Una reorganización de bancarrota hace esto. Dado que la compañía continuará operando, es muy natural que se permita a la administración actual, antes que a un administrador fiduciario, dirigirla durante el periodo de reorganización, y que se deje en manos de esa administración la iniciativa de la reorganización. Los miembros de la administración tienen experiencia relevante y un fuerte incentivo para reorganizarse con éxito a fin de salvar sus empleos y evitar así la pérdida del capital humano específico de la empresa.

Pero, una vez dicho todo esto, subsiste la interrogante de por qué, si la empresa vale más como empresa solvente que cerrada, los acreedores no consentirían en una reorganización voluntariamente. ¿Por qué debiera la ley permitirle (como lo hace) a un tribunal imponer un plan de reorganización de la bancarrota sobre una minoría de acreedores que no están de acuerdo? Hay dos posibles respuestas. Una es conocida: el problema del gorroneo. Si se requiriera el consentimiento unánime de los acreedores para que se aprobara la reorganización, le convendría a cada acreedor rehusarse en busca de un tratamiento preferente en la asignación de las acciones en la entidad reorganizada. La otra respuesta es más interesante: es posible que la reorganización sea lo más conveniente para los accionistas y los acreedores en conjunto, pero que perjudique a algunos acreedores. Según la regla de la prioridad absoluta (véase el problema 4 al final del capítulo siguiente), un acreedor más antiguo tiene derecho a que se le pague todo antes de que los más recientes reciban nada, de modo que se le podría asegurar el pago completo si la empresa es liquidada, y no quisiera intercambiar este aseguramiento por sus perspectivas como accionista en la empresa reorganizada. Tiene derecho a recibir en la reorganización un interés accionario equivalente, en el rendimiento esperado ajustado por el riesgo, a lo que recibiría en la liquidación. Pero la equivalencia es difícil de calcular por diversas razones; el valor de un interés accionario depende de las perspectivas de ganancias de la empresa, las que podrían ser muy inciertas. Además, la regla del "valor nuevo" de la Ley de Quiebras faculta a la persona en bancarrota a otorgar una primera prioridad a los prestamistas que le otorguen crédito nuevo a la empresa (de otro modo sería difícil inducir a los prestamistas a prestar dinero a las personas en bancarrota). Esto podría debilitar las perspectivas de pago completo del acreedor antiguo, en el caso de que la empresa fuera liquidada. Cuando la liquidación inmediata produce menos dinero para pagar a los acreedores que una continuación del negocio reorganizado, los costos se imponen a los acreedores. Pero un acreedor que a causa de su antigüedad tenga más que ganar de la liquidación que de la reorganización no considerará esto al decidir si aceptará o no la reorganización. Aunque en principio podría ser compensado este acreedor por los otros acreedores para que acepte, los costos de la tran-

sacción necesaria de múltiples partes podrían ser formidables incluso sin considerar los problemas del gorroneo.

Así pues, para los acreedores plenamente asegurados, una venta de remate de las acciones de la empresa puede ser preferible a una reorganización aun cuando la empresa posea más valor siendo reorganizada que si se la liquidara. Éste fue un problema en la bancarrota de Chrysler en 2009. La sección 363 del Código de Quiebras autoriza al juez de bancarrotas a subastar las acciones de la empresa en quiebra si ése es el mejor trato para los acreedores; los ingresos de la subasta se distribuyen entre los acreedores de acuerdo con las normas de prioridad acostumbradas. Las ventas de la sección 363 se han vuelto una alternativa común a las reorganizaciones corporativas en la bancarrota.[23] El primer pujador, que en el caso de la quiebra de Chrysler fue el gobierno federal, le propone al juez de bancarrotas la forma de estructurar las pujas. El gobierno propuso que a cada pujador se le impusiera el requisito de ofrecer tanto como (el propio gobierno) estuviera dispuesto a ofrecer a los planes de retiro del sindicato. El juez estuvo de acuerdo. La consecuencia de esto fue que se hacía improbable que la subasta atrajera pujas que honraran la prioridad de los acreedores asegurados, y de hecho no logró hacerlo. Para ilustrar el punto, supongamos que los acreedores antiguos tienen reclamos de 4 mil millones y que la estructuración de la puja propuesta por el gobierno y aceptada por el juez requiere que un pujador cualquiera prometa 3 mil millones en beneficio de los planes de retiro del sindicato. Supongamos, más aún, que la puja del gobierno es de 5 mil millones, de manera que los acreedores antiguos recibirían solamente 2 mil millones. Es de concebirse que algún otro postor interesado podría haber estado dispuesto a pagarles el total de los reclamos a los acreedores antiguos pero indispuesto a efectuar a un tiempo ese pago *y* aportar nada menos que otros jugosos 3 mil millones de dólares para los planes de retiro; pues demos por sentado que él evaluaba las acciones de la compañía en solamente 4 mil millones. La puja de ese postor, que restauraría a los acreedores asegurados a su condición inicial, no sería elegible.

Estos acreedores asegurados sostuvieron infructuosamente que la estratagema del gobierno violaba la norma de prioridad absoluta. Pero, de ser así, la violación puede haber estado económicamente justificada. Supóngase que los acreedores antiguos en nuestro ejemplo numérico, no estando convencidos de que la empresa tenga futuro alguno como empresa solvente, ofrecen 4 mil millones por las acciones con el propósito de venderlas fragmentadas, liquidando de esta forma a la compañía. El gobierno, al no estar dispuesto a

[23] Véase Douglas G. Baird, "Lessons from the Automobile Reorganizations", 4 *Journal of Legal Analysis*, 271 (2012). A menudo estas ventas son arregladas o forzadas por sus acreedores antiguos, quienes no quieren que la administración actual de la empresa siga ejerciendo el control; en realidad, con frecuencia se opta por la reorganización prefiriéndola a la liquidación a fin de proteger tanto tiempo como sea posible, retrasando lo inevitable, los salarios y las ventajas adicionales de la administración.

pagar más de 5 mil millones y, no obstante, pensar que sería necesario realizar un pago de 3 mil millones para los planes de retiro a fin de dar a la compañía una oportunidad razonable de sobrevivir como empresa solvente, no se ajustaría a la puja de dichos acreedores si éstos poseyeran el derecho a condiciones de venta que les reportaran un pago antes que a todos y cada uno de los acreedores no asegurados, tales como los planes de retiro. Eso constituiría un resultado poco atractivo, ya que arrojaría un precio inferior para las acciones: 4 mil millones en vez de 5 mil millones de dólares. Una compañía que, según el punto de vista de un postor, valiera más como empresa solvente que yéndose a la liquidación sería vendida a una puja inferior. Un objeto central del derecho de bancarrotas consiste en maximizar el valor del bien inmueble del deudor, y a veces ese propósito se puede consumar sólo a expensas de los acreedores asegurados.

En el caso de Chrysler, la reorganización (en términos técnicos una venta de acciones, pero desde el ángulo funcional una reorganización porque la compañía permanecía en la esfera mercantil) preservó el valor; no obstante, a menudo la liquidación podría preservar más valor que la reorganización pero enfrenta resistencia, y ello no sólo debido a que con ella los administradores habrán de perder su trabajo y, por consiguiente, su capital específico de la empresa; acreedores nuevos con escasas probabilidades de recuperar siquiera algo en una liquidación exhiben su interés en mantener viva a la empresa aun cuando sus propias acciones tengan más valor en una liquidación. (También a los accionistas les agrada el capítulo 11, siempre que la reorganización les otorgue por lo menos un pequeño interés accionario en la empresa reorganizada. Porque entonces la reorganización es para ellos una proposición de no perder. Si la empresa reorganizada es rentable, compartirán las ganancias; si se derrumba, toda la pérdida recaerá en los acreedores.) Pero también es cierto que, si no hubiera un capítulo 11, los administradores lucharían con todas sus fuerzas para evitar la quiebra, porque la quiebra significaría la liquidación.

La balanza entre los administradores y los acreedores antiguos podría inclinarse demasiado a favor de los primeros debido a una asimetría del poder de negociación generada por las reglas de la Ley de Quiebras. Durante el periodo en el que los administradores tienen el derecho exclusivo de proponer un plan de reorganización, y mientras el plan está bajo la consideración del tribunal de quiebras, los acreedores asegurados no pueden hacer efectiva la garantía, ni los otros acreedores pueden forzar una liquidación. En suma, los acreedores carecen de una "opción de salida" negociando un plan de reorganización. Los administradores sí tienen tal opción; están en libertad para renunciar a la empresa y buscar trabajo en otra parte. Entre más valiosa sea esa opción, mejores serán los términos que exigirán a los acreedores.[24]

[24] Douglas G. Baird, Robert H. Gertner y Randal C. Picker, *Game Theory and the Law,* 232-237 (1994). El resultado son desviaciones frecuentes de la regla de prioridad absoluta. Maria

Otro problema es lo que se ha llamado "falla de filtración".[25] Hay dos tipos de corporaciones en bancarrota: *1)* las que son viables y debieran continuar en operación y *2)* las que no son viables y debieran ser liquidadas. Los administradores de ambos tipos prefieren la reorganización, por la razón sugerida antes. Los administradores de una empresa inviable tienen un interés obvio en hacerla aparecer como viable, para que los acreedores no obliguen a liquidarla. Menos obviamente, los administradores de una empresa viable tienen interés en hacerla aparecer menos viable de lo que es, porque ello llevará a los acreedores a exigir menos en el plan de reorganización, lo que a su vez aumentará los beneficios para los accionistas, con quienes los administradores están ligados. A resultas de estos incentivos, los acreedores y el tribunal de la quiebra tendrán dificultades para decidir cuáles corporaciones en quiebra debieran ser preservadas y cuáles disueltas. Este equilibrio combinado podría ser la razón de que la mayoría de las reorganizaciones por quiebra terminen en la liquidación, lo que implica que el efecto del capítulo 11 es retrasar la reasignación de los activos de la empresa en bancarrota más allá del punto óptimo.[26]

De manera creciente, los acreedores de las compañías en el caso de la bancarrota del capítulo 11 están empleando el financiamiento (préstamo tomado por la empresa en quiebra con la aprobación del juez de bancarrotas) de la modalidad "deudor en posesión" para guiar la reorganización de tal manera que limite el control ejercido por los administradores de las compañías.[27] La deuda debida al prestamista recibe prioridad sobre todos los acreedores del deudor previamente a la quiebra. El financiamiento de deudor en posesión no es ninguna novedad. Lo que resulta novedoso es que sus prestamistas se propulsan hacia la gobernanza de la empresa, tanto al ejercer presión sobre la junta directiva (algo que está al alcance de los acreedores porque comúnmente terminan siendo los accionistas de la empresa reorganizada cuando ésta sale de la bancarrota, y sabiendo esto los directivos cooperarán con ellos desde el inicio) como al motivar a los administradores con contratos de "remuneración basada en el desempeño". Con todo, preocupa que cuando el acreedor del deudor en posesión es también un acreedor previo a la petición pueda ejercer su influencia sobre la administración de la empresa con el fin de favorecer de su deuda anterior a la petición.

La quiebra corporativa se ha vuelto más frecuente a partir y quizás a causa de la invención de la adquisición apalancada: una transacción en la

Carapeto, "Explaining Deviations from Absolute Priority Rules in Bankruptcy", 3 *J. Empirical Legal Stud.,* 543 (2006).

[25] Michelle J. White, "Corporate Bankruptcy as a Filtering Device: Chapter 11 Reorganizations and Out-of-Court Debt Restructurings", 10 *J. Law. Econ. & Organization,* 268 (1994).

[26] Véase Douglas G. Baird y Edward R. Morrison, "Bankruptcy Decision Making", 17 *J. Law, Econ. & Organization*, 356 (2001). La alternativa que sugieren es la subasta obligatoria de la firma en quiebra.

[27] Véase David A. Skeel, Jr., "Creditors' Ball: The 'New' New Corporate Governance in Chapter 11", 152 *U. Pa. L. Rev., 917* (2003).

que se compra una compañía en gran medida con dinero que se toma prestado y el prestatario-comprador asegura el préstamo con los propios activos de la compañía. Por ejemplo, los administradores de una empresa podrían comprarla a sus dueños con dinero obtenido en préstamo de un banco al que los administradores ofrecen los activos de la empresa como garantía. El resultado es un aumento de la cantidad de deuda en los libros de la empresa y, por ende, en el riesgo de una quiebra. ¿Es esto algo malo? El trueque entre ventajas y desventajas es complejo.[28] Por una parte, en la medida en que la quiebra impone pérdidas sociales de peso muerto —es decir, no sólo causa una transferencia de riqueza de los accionistas, los administradores y algunos acreedores a otros acreedores, sino también significa el consumo de recursos valiosos (el tiempo de abogados, banqueros y jueces, más allá de lo que se consumiría en el litigio de cobranza ordinario), así como las reducciones de la eficiencia en el uso de los activos que discutimos antes—, cualquier cosa que aumente el riesgo de la quiebra impone un costo social. Por otra parte, el aumento de la cantidad de deuda de una empresa permite que los administradores se vuelvan accionistas controladores, lo que reduce los costos de agencia dentro de la corporación al convertir a los agentes en principales. Esto es más probable que ocurra cuando la corporación es de propiedad pública y los administradores tienen un interés accionario bajo antes de la transacción. En una corporación pública los administradores, si no son accionistas de envergadura, pueden usar sus facultades de control para gestionar a la compañía de tal modo que sirva para incrementar su riqueza personal en lugar de maximizar las ganancias corporativas. Este conflicto de intereses se elimina al convertir a los administradores en dueños por medio de una adquisición apalancada.

Una adquisición apalancada puede reducir los costos de agencia también indirectamente, ya que disminuye la capacidad de la empresa para financiar proyectos nuevos con las ganancias retenidas y así la obliga a acudir al mercado de capital con mayor frecuencia, reduciendo de este modo el peso de la discreción de los administradores (véase después XV.14). Y, al convertir a la empresa en una aventura más riesgosa, una razón de deuda-capital accionario elevada podría contrarrestar la tendencia de los administradores a comportarse de un modo contrario al riesgo. Los accionistas desean de ordinario que las corporaciones en las que invierten se comporten como si fueran neutrales al riesgo, porque el accionista adverso al riesgo puede protegerse contra el riesgo específico de la empresa simplemente manteniendo una cartera de valores diversificada. Pero el capital humano específico de la empresa

[28] Véanse, por ejemplo, Bengt Holmstrom y Steven N. Kaplan, "Corporate Governance and Merger Activity in the United States: Making Sense of the 1980s and 1990s", *J. Econ. Perspectives,* primavera de 2001, pp. 121, 128; Steven N. Kaplan y Jeremy C. Stein, "The Evolution of Buyout Pricing and Financial Structure in the 1980s", 108 *Q. J. Econ.,* 313, 340-344 (1993); Douglas G. Baird, "Fraudulent Conveyances Agency Costs, and Leveraged Buyouts", 20 *J. Legal Stud.,* 1, 8-9 (1991).

de los administradores no puede diversificarse fácilmente, de modo que los administradores podrían tener una tendencia a administrar las corporaciones de un modo adverso al riesgo. Esta tendencia puede contrarrestarse mediante el apalancamiento, el que magnificará el efecto financiero de las fluctuaciones del desempeño de la empresa. (En el capítulo siguiente volveremos a ocuparnos de este tema.) Con todo, hay una complicación cuando los compradores de la adquisición apalancada son inversionistas antes que los administradores. El mayor riesgo de fracaso financiero generado por la razón mayor de deuda-capital accionario resultante de la adquisición podría volver a los administradores más tímidos aún, al mismo tiempo que vuelve más riesgosas las consecuencias de sus decisiones gerenciales.

El abuso del instrumento de la adquisición apalancada se ilustra con un caso donde el peso de la deuda creada por la transacción era tan oneroso que la empresa no tenía ninguna perspectiva razonable de sobrevivir.[29] El efecto de la transacción consistía nada más en borrar de un plumazo cualesquiera posibilidades de que los acreedores no asegurados recobraran una mínima parte de sus reclamaciones. El pago recibido por los accionistas de manos del comprador de la corporación fue por ende tachado, en una demanda levantada por el administrador fiduciario por bancarrota (en representación de los acreedores no asegurados), como un traspaso fraudulento. Pues, a cambio del efectivo que los accionistas habían recibido (dinero que el comprador obtuvo al imponer una carga sobre los activos de la firma, ya abrumados por cargas), no habían aportado valor alguno a la corporación; la transacción simplemente aumentó la deuda de la corporación y, al hacerlo, la empujó al precipicio.

XV.6. La quiebra personal

Ahora tenemos que considerar por qué la ley permite que haya quiebras no mercantiles voluntarias (por ejemplo, de los consumidores que se pasan de la raya al comprar artículos a crédito). Un número de casos dan cuenta de la postura de dar a la víctima de la quiebra la oportunidad de un "borrón y cuenta nueva", si bien esa óptica sigue una lógica más aplicable al tratamiento de las quiebras corporativas. Aun así, muchos pequeños negocios no se conducen a la manera de las corporaciones; y una deuda aplastante puede hacerle difícil a una persona conseguir y mantener un empleo, o hacerse de la educación que requiere para obtener un trabajo. Así pues, la lógica del "borrón y cuenta nueva" tiene después de todo aplicación tanto en el caso de las quiebras corporativas como en el de las individuales. Y, por añadidura, hay otras respuestas a la pregunta sobre el motivo por el que la ley debiera ofrecer a los individuos la oportunidad de partir de cero. Una respuesta, basada en la

[29] Boyer *vs.* Crown Stock Distribution, Inc., 587 F.3d, 787 (7th Cir., 2009).

aversión al riesgo, es que la bancarrota proporciona una especie de seguro, si bien no un seguro de negocios, que es difícil de comprar en el mercado. No puede uno comprar un seguro privado contra la eventualidad de quebrar, ya que entonces la gente realmente pediría préstamos en forma desenfrenada; pero incluso un prestatario prudente se podría encontrar sumido en la bancarrota debido a una racha de mala suerte. Y, si a los deudores se les prohibiera desprenderse de sus deudas declarándose en bancarrota, los tribunales incurrirían en costos gravosos al supervisar la recaudación de las deudas que a lo largo de su existencia contrajo la persona arruinada. Esto puede explicar por qué no se le permite a una persona, cuando pide prestado dinero, renunciar a su derecho de buscar descargo de deudas en la bancarrota. Los costos que representa para el sistema judicial supervisar la recolección de deudas a lo largo de un periodo largo (la alternativa al descargo es el pago de acuerdo con un plan de abonos por un tiempo indefinido) serían considerables y no serían soportados en su totalidad por las partes involucradas en el préstamo.

La objeción planteada a la práctica de dar espacio a las quiebras personales es equivalente a la razón por la que es difícil comprar un seguro contra el impago de las deudas de uno: le da al impago un cariz atractivo. Uno puede comprar y consumir todo tipo de cosas agradables a crédito y a continuación incumplir. Este problema, que es una forma particularmente grave del problema general del azar o riesgo moral que reviste el seguro, podría en principio resolverse al distinguir entre impagos voluntarios e involuntarios y reservar el privilegio de la quiebra voluntaria a los últimos. Empero, sería difícil hacer esta distinción en la práctica. No existe una demarcación nítida entre los impagos voluntarios y los involuntarios. Entre más deudas contrae uno, más probable es que sea incapaz de pagar en virtud de circunstancias imprevistas, sin importar cuán genuinas puedan ser las intenciones de saldarlas.

Repárese en el efecto que tendría en la redistribución de la riqueza una Ley de Quiebras personales de sello liberal si los costos de información les impidieran a los prestamistas ajustar sus tasas de interés de acuerdo con la probabilidad de que el prestatario se declare en bancarrota. El efecto habrá de ser el de hacer que los prestatarios prudentes subsidien a los precipitados… ¡una curiosa forma de apuntalar la redistribución de la riqueza! Pues, al contrario de lo que uno pudiera suponer, una Ley de Quiebras de cuño liberal no redistribuye la riqueza haciéndola desplazarse de los acreedores en cuanto clase a los deudores en cuanto clase. Sólo determina que los acreedores eleven sus tasas de interés a fin de contrarrestar el riesgo añadido de impago que dicha ley genera. Sin embargo, si la ley no es anticipada y se aplica retroactivamente, esto es, a los contratos crediticios en vigor, habrá por una sola vez una redistribución a favor de los deudores existentes y en contra de los acreedores existentes.[30]

[30] Como aparentemente sucedió a resultas de los cambios realizados por el Código de Quiebras de 1978. William J. Boyes y Roger L. Faith, "Some Effects of the Bankruptcy Reform Act of 1978", 29 *J. Law & Econ.,* 139 (1986).

La Ley de Prevención del Abuso y Protección al Consumidor de 2005 ("Ley de Reforma de la Quiebra" para abreviar) hizo más difícil a los individuos declararse en quiebra según el capítulo 7 del Código de Quiebras. De acuerdo con el capítulo 7, los activos de quien sufre quiebra, menos los activos eximidos, como una casa y las herramientas de trabajo, se venden para pagar a los acreedores. Cuando quien cae en quiebra es un sujeto y no una corporación, generalmente sus activos son demasiado magros para que los acreedores puedan recuperar todo lo que se les debe; con frecuencia reciben sólo algunos centavos de un dólar. Un procedimiento alternativo al que pueden acogerse los individuos (y sus acreedores) es el capítulo 13 sobre la bancarrota, de acuerdo con el cual, en vez de ceder sus activos no eximidos a los acreedores, el deudor acepta realizar pagos periódicos a nombre de ellos durante un periodo de no menos de cinco años después de la declaración de quiebra. Aunque el capítulo 13 es atractivo para quienes poseen activos en cantidad sustancial, el capítulo 7 lo es para aquellos individuos que tienen pocos activos no eximidos, esto es, para la mayoría de los individuos. El capítulo 7 del Código los faculta para acumular enormes deudas con tarjeta de crédito por compras que no crean activos duraderos (como comida, viajes y entretenimiento), declararse en quiebra, saldar sus deudas y luego empezar de nueva cuenta. (Aun cuando existen límites sobre cuán a menudo puede una persona declararse en bancarrota, no hay límite alguno a la bancarrota involuntaria.) Las quiebras personales se habían vuelto estratosféricas debido a que se había expandido la emisión de tarjetas de crédito, lo que aumentó el acceso al crédito, junto con la prevención federal de las leyes sobre la usura, lo que dio paso al préstamo a tasas de interés muy altas, las cuales a un tiempo reflejan e incrementan el riesgo de impago. La Ley de Reforma de la Quiebra obliga a muchos deudores que tienen ingresos que exceden al ingreso medio en su estado de residencia a optar por la ruta del capítulo 13 y, por lo tanto, a realizar pagos provenientes de su ingreso durante cierto número de años. La ley igualmente extiende el periodo durante el que un sujeto en quiebra debe aguardar, una vez que ha sido liberado de sus deudas existentes, antes de que pueda declararse nuevamente en quiebra y deshacerse de un nuevo ciclo de deudas.

Los críticos de la reforma sostienen que más de la mitad de todas las quiebras no corresponden a prestatarios desaforados: se trata de gente desventurada que ha sufrido el embate de gastos médicos inesperados. Pero que uno se vea forzado a irse a quiebra como consecuencia de gastos médicos (o por haber perdido el empleo a resultas de un problema médico) puede derivar de que haya tomado una serie de préstamos anteriormente. Si uno ya está endeudado hasta el cuello por haber pedido préstamos, un gasto médico súbito puede de hecho empujarlo al abismo. Ahora bien, sabiendo que los gastos médicos son un riesgo en nuestra sociedad, las personas juiciosas evitan sobrecargarse hasta el tope con deudas no médicas.

El verdadero problema con la reforma es que, en la medida en que su propósito es reducir el sobreendeudamiento del consumidor (un nivel de deuda que crea un elevado riesgo de impago por el carácter fijo de ésta), puede resultar fallida porque al aumentar los derechos de los prestamistas reduce las tasas de interés, dado que éstas son en parte una función del riesgo de impago.

Una interrogante cuya respuesta cobra urgencia a consecuencias del colapso en el mercado inmobiliario —que cuenta decisivamente entre los precipitantes del colapso financiero de septiembre de 2008, el cual a su vez precipitó la recesión económica de la que los Estados Unidos todavía se están recuperando— es si puede o debería usarse la quiebra para permitirles a los propietarios de una casa efectuar el despojo de hipotecas que con toda probabilidad jamás podrán ser liquidadas porque sobrepasan en un grado excesivo el valor de la propiedad hipotecada. Recordemos la máxima de que los bienes retenidos como garantía de pago, al atravesar el proceso de quiebra, salen ilesos. Un banco que tiene una hipoteca que supera en grado mayúsculo el valor actual de la propiedad hipotecada puede no molestarse en levantar un reclamo sobre la quiebra del deudor hipotecario o en dar cierre a la hipoteca, pues en ninguno de estos casos tendrá la expectativa de recobrar mucho dinero, y al cabo de cualquiera de esos procedimientos habrá tenido que incurrir en considerables gastos legales. Será especialmente remota la posibilidad de que el banco apele a cualquiera de estos remedios si es un segundo acreedor hipotecario y el valor de la propiedad hipotecada es considerablemente menor al de la primera hipoteca. Puede preferir no hacer nada hasta que, y a menos que, el valor de la propiedad se incremente repentina y notablemente.

Hay en el capítulo 7 del Código de Quiebras provisiones referentes a la forma de atenuar o echar abajo el derecho de retención de bienes. La sección 522(f) le permite al deudor reducir un derecho de retención sobre la propiedad exentada todo lo que sea necesario para preservar la exención, mientras que la sección 722 le permite redimir "propiedad personal tangible destinada fundamentalmente al uso personal, familiar u hogareño, de un derecho de retención de bienes como garantía de pago que asegure la descarga de una deuda liquidable del consumidor" mediante el pago del valor actual del bien retenido. Ambas disposiciones respaldan la política de "borrón y cuenta nueva" del capítulo 7. No obstante, considérese ahora este caso reciente.[31] Un banco tenía una segunda hipoteca sobre una casa por 50 000 dólares, mientras que otro prestamista tenía una primera hipoteca sobre esa casa por 243 000 dólares… aun cuando el valor de la casa era en realidad de sólo 165 000 dólares. Quebrados, los dueños del inmueble deseaban que el tribunal de quiebras dictara la invalidación (el despojo) de la segunda hipoteca, con base en la teoría de que carecía de valor. De hecho, tenía un valor prospectivo; algún día la casa podría tener un valor mayor al de la primera hipoteca, y entonces

[31] Palomar *vs.* First American Bank, 2013 WL, 3466884 (7th Cir., 11 de julio, 2013).

el segundo acreedor hipotecario podría cerrar el proceso con una expectativa de recuperar parte del dinero que se le debía. No había forma de poner en duda la validez de la segunda hipoteca.

El haber arrumbado la segunda hipoteca no habría promovido la política de "partir desde cero", por lo que no es de sorprender que el tribunal hallara que el capítulo 7 no autorizaba semejante despojo. Todo lo que hacía la existencia ininterrumpida de la segunda hipoteca era privar a los deudores de la oportunidad de hacerse de algo de dinero en el caso de que el valor de su casa excediera alguna vez el saldo respecto a la primera hipoteca. Privar a los deudores de una oportunidad futura de carácter a tal punto especulativo no podía concebirse realistamente como algo que cancelara su habilidad para volver a empezar desde cero. Al invalidar el derecho de retención no se les facultaría para obtener una nueva y segunda hipoteca (a menos que proviniera de un prestamista depredador) ni de otro modo se ayudaría a que mejorara su situación financiera.

XV.7. Tras el velo corporativo

Aunque la responsabilidad limitada desempeña una función importante para volver más atractivas las inversiones en capital para los individuos, hacer caso omiso de ella —mirar detrás del velo corporativo, en la jerga del derecho de las corporaciones— podría promover la eficiencia en dos situaciones. La primera y más problemática es el caso de la subcapitalización, ilustrada por la compañía de taxis cuyos propietarios incorporan cada taxi por separado a fin de limitar la responsabilidad cuasidelictiva ante las víctimas de accidentes, con lo que se externaliza un costo del servicio de taxis.[32] Pero es posible que la penetración del velo no resulte el remedio ideal para esta situación. Si se permitiera que las víctimas de cuasidelitos se protegieran con los activos de los accionistas, se generaría una fuente de riesgo adicional para los accionistas, y un aumento del riesgo, como sabemos, es un costo real para quienes sienten aversión por el riesgo. La compañía podría asegurarse contra sus cuasidelitos, pero ésta no sería una alternativa satisfactoria, en especial cuando el riesgo de responsabilidad es masivo, como en el caso de los accidentes del reactor nuclear o en el de la enfermedad pulmonar provocada por el asbesto. Y desgarrar el velo es un escenario de pesadilla en el ámbito administrativo cuando hay muchos accionistas y con frecuencia se pasa revista a las acciones. Pues entonces uno debe decidir cuáles accionistas habrán de ser respon-

[32] Se aboga por mirar detrás del velo corporativo en Henry Hansmann y Reinier Kraakman, "Toward Unlimited Shareholder Liability for Corporate Torts", 100 *Yale L. J.,* 1879 (1991). Adviértase que, si el velo es rasgado, debiera exigírseles a las corporaciones no afiliadas de servicio de taxi sólo de ida que peguen un cartel para impedir que externalicen sus costos de accidentes, acción mediante la que obtienen una ventaja competitiva que es ineficiente.

sables —¿aquellos que poseían acciones cuando ocurrió el cuasidelito, aquellos que las poseían cuando tuvo lugar el juicio por cuasidelito o quienes las poseían en algún otro momento?— y cómo se repartirá la responsabilidad entre los accionistas del grupo responsable.

Una alternativa a perforar el velo corporativo consiste en requerir a cualquier corporación dedicada a realizar una actividad peligrosa que otorgue una fianza igual a la estimación razonable más elevada de la probable cuantía de su responsabilidad cuasidelictiva. Los accionistas quedarían protegidos contra el aniquilamiento total, aunque sus ganancias se reducirían. La cuestión espinosa es ¿quién determina el monto de la fianza? Tendría que ser una dependencia gubernamental dotada de algún modo de un conocimiento amplio de las potenciales responsabilidades cuasidelictivas en cada área de la economía.

El argumento menos problemático a favor de perforar el velo corporativo se aplica cuando la incorporación separada engaña a los acreedores. Cuando se permite que las corporaciones hagan creer que tienen activos para pagar a sus acreedores en mayor cuantía de los que en realidad poseen, aumentan efectivamente los costos en los que deben incurrir los acreedores para determinar la verdadera capacidad crediticia de las corporaciones con las que tratan.

Al decidir si perforarán el velo corporativo, los tribunales son dados a preguntar si la corporación deudora es simplemente un agente, un *alter ego* o un instrumento del accionista y, por lo tanto, consideran si el deudor ha observado las formalidades corporativas (por ejemplo, si regularmente ha celebrado reuniones de directores) y si se encuentra adecuadamente capitalizado. Estos criterios carecen de utilidad. La prueba de *alter ego* no contesta la pregunta, el carácter instrumental o de agencia de la corporación es irrelevante, la observancia de las formalidades corporativas es una consideración trivial y la preocupación por la capitalización deficiente es perversa, muy aparte de los problemas derivados de este fenómeno y de los que nos ocupamos al poner el ejemplo de la compañía de taxis. Las empresas que empiezan sus operaciones están de ordinario escasamente capitalizadas, pero son precisamente aquellas cuyos accionistas necesitan más la responsabilidad limitada como un incentivo para asumir el riesgo empresarial, porque una gran parte de las empresas nuevas fracasan y porque sus inversionistas son a menudo personas de recursos modestos. Quienes prestan a corporaciones nuevas o escasamente capitalizadas pueden exigir que los accionistas acepten garantizar las deudas de la corporación; si no lo hacen, la ley no debiera sacarles las castañas del fuego a menos que se haya presentado engañosamente la capacidad financiera de la corporación.

A menudo, un accionista es una corporación antes que un individuo, y podría parecer que la política de desplazamiento del riesgo que se encuentra detrás del principio de la responsabilidad limitada no se aplicaría en ese caso. Si se hace responsable a una corporación matriz por las deudas de su subsidiaria, la exposición de los accionistas de la matriz a la responsabilidad, aun siendo

mayor que si la subsidiaria disfrutara de una responsabilidad limitada, está limitada a su inversión en la matriz y puede reducirse más aún por su tenencia de una cartera de acciones diversificada. Sin embargo, es necesario distinguir entre la corporación de acciones públicas (muchos accionistas, acciones regularmente negociadas) y la corporación cerrada (pocos accionistas, fuera del mercado de valores). Supongamos que el señor A. Smith desea invertir en una empresa minera, pero toda la fortuna de Smith (aparte de la que Smith planea comprometer en el proyecto minero) está invertida en una estación de radio propiedad de una corporación de la que Smith es el único accionista. Si Smith forma una nueva corporación para llevar a cabo el proyecto minero, y si los activos de las corporaciones afiliadas pueden ser reunidos para satisfacer las reclamaciones de los acreedores de una de las filiales, Smith habrá apostado toda su fortuna al resultado del proyecto minero. En este caso no hay ninguna diferencia entre perforar el velo corporativo para llegar a los activos de una corporación afiliada y hacerlo para llegar a los activos de un accionista individual.

Incluso cuando una corporación grande, pública, opera a través de subsidiarias cuya propiedad total le pertenece, mientras esas subsidiarias se encuentren en negocios no relacionados, la maximización de las ganancias de la corporación matriz requerirá que las ganancias de cada subsidiaria se maximicen por separado, de modo que los activos, los costos y las deudas de cada subsidiaria deberán ser los mismos que en el caso de que fuesen empresas separadas; de este modo, los acreedores no se ven perjudicados por estar limitados a sus derechos contra la subsidiaria, en particular con la que han tratado. El peligro del abuso de la forma corporativa es mayor en el caso de la empresa pequeña, donde la operación de las corporaciones que la constituyen como centros de ganancias separados es menos necesaria para asegurar la administración eficiente, pero donde el interés de los inversionistas individuales por la responsabilidad limitada de las filiales corporativas se aproxima al interés del inversionista por preservar la responsabilidad limitada de las corporaciones no afiliadas. Ése es nuestro ejemplo del señor Smith.

Aun cuando las actividades de las corporaciones afiliadas estén estrechamente relacionadas —cuando, por ejemplo, producen bienes complementarios—, cada corporación será operada normalmente como un centro de ganancias separado.[33] Es cierto que si hay ahorros sustanciales de costos por la propiedad común, como en algunos casos donde las corporaciones afiliadas operan en etapas sucesivas de la producción de un bien (véanse antes X.10,

[33] No obstante, si una compañía que vende un producto adquiriera otra que vende un producto diferente y, luego, atara el segundo al primero como un método de discriminación de precio (según se describe en el capítulo sobre los antimonopolios), tendería a fijar un precio más bajo para el precio atador y uno más alto para el producto atado que si las empresas fueran de propiedades por separado. Por consiguiente, las ganancias para las compañías serían diferentes al caso de que fueran propiedades independientes una de la otra. Véase antes X.10.

XV.1), las dos corporaciones se administrarán de manera diferente de las corporaciones de propiedad separada en la misma línea de negocios; sus operaciones estarán integradas de manera más estrecha que las de corporaciones independientes. Pero sería perverso castigar a tal corporación por su eficiencia superior retirándole el privilegio de la responsabilidad limitada de la que disfrutan sus competidores no integrados. Además, el propietario común tiene un fuerte incentivo para evitar las transferencias intercorporativas que, al distorsionar la rentabilidad de cada corporación, le vuelven más difícil la evaluación de su desempeño. Es por ello que el precio al que una división de una empresa verticalmente integrada "venderá" su producción a otra división es normalmente el precio de mercado del bien en cuestión (menos cualesquiera ahorros de costos imputables a la realización de una transferencia dentro de la empresa por comparación con una transacción de mercado), antes que un precio de transferencia arbitrario destinado a aumentar las ganancias de una división a expensas de la otra.

La diferencia importante entre un grupo de filiales que opera en negocios relacionados y un grupo que opera en varios negocios no relacionados es que en el primer caso el acreedor tiene mayor probabilidad de engañarse creyendo que está tratando con una sola corporación. Por ejemplo, supongamos que una compañía controladora bancaria establece una subsidiaria para invertir en inmuebles. La compañía controladora le pone a la subsidiaria un nombre equívocamente similar al de la subsidiaria bancaria de la compañía controladora y la corporación inmobiliaria renta espacio de oficinas en el banco, de modo que sus oficinas parecen ser oficinas bancarias. Los acreedores poco sofisticados otorgan tasas de interés bajas a la subsidiaria inmobiliaria creyendo razonablemente que están tratando con el banco mismo. En estas circunstancias podría ser conveniente que se prohibiera a la compañía controladora bancaria —o aun al banco mismo— negar ser la entidad a la que los acreedores han otorgado crédito. Sin embargo, adviértase que el principio de la presentación engañosa es adecuado para el manejo de estos casos. Empero, adviértase que, cuando no hay una presentación engañosa, una regla que abrogue la responsabilidad limitada de las corporaciones afiliadas no reduciría los riesgos de ninguna clase de acreedores: sencillamente aumentaría sus costos de información. Aunque el acreedor de la corporación A supiera que si A no pagaba podría él alcanzar los activos de su afiliada B, también sabría que si B dejara de pagar los acreedores de B podrían tener un derecho sobre los activos de A, lo que podría hacer que ésta dejara de pagar la deuda que tiene con él. Por lo tanto, para saber cuán elevada es la tasa de interés que debe cobrar, este acreedor tendría que investigar la situación financiera de B al igual que la de A. Y B podría estar en un negocio nada relacionado con el de A.

XV.8. La separación de la propiedad y el control

Generalmente se piensa que el derecho de las corporaciones no está orientado sólo en contra de los acreedores de las corporaciones sino también en contra de los propios accionistas. (Estas creencias no son por necesidad incongruentes; si la ley está inclinada contra los acreedores, lo que trae por consecuencia costos de interés más elevados para la corporación, los accionistas habrán de sufrir.) La razón de esto, según se aduce, es la separación entre la propiedad (los accionistas) y el control (los administradores) en la gran corporación pública. Gran parte de la preocupación por esta separación se basa en una falta de distinción entre la *empresa* y la *corporación* (véase antes XV.1). La empresa es un método de organización de la producción; la corporación es un método, como el otorgamiento de una fianza, para atraer capital hacia la empresa. El negocio grande típico es a la vez una empresa y una corporación. El control de la empresa reside en un grupo de administradores que gira órdenes a los empleados que compran los insumos de la empresa y elaboran y venden su producto. El grupo administrador está integrado por personas que tienen experiencia en el negocio y participan en él a tiempo completo, día tras día. El accionista típico (excepto en la corporación cerrada o donde un accionista es propietario de un porcentaje muy grande de las acciones de la corporación) no conoce el negocio de la empresa, no obtiene de ella una parte importante de sus ingresos, y ni espera ni tiene un incentivo para participar en su administración. Es un inversionista pasivo y, en virtud de la liquidez de su interés, tiene con la empresa sólo una relación casual y a menudo transitoria. Su interés, como el de un acreedor, es un interés financiero antes que administrativo.

El hecho de que los accionistas no administren ni controlen "su" corporación no es más anómalo que el hecho de que los tenedores de bonos no administren ni controlen las corporaciones, o los beneficiarios de un fideicomiso al administrador fiduciario. Los tres grupos tienen un interés de inversión. La diferencia reside en la mayor vulnerabilidad del accionista (como la del beneficiario de un fideicomiso) que la del tenedor de un bono ante la mala actuación o la falta de actuación de los administradores corporativos. Dado que el tenedor de un bono tiene una tasa de interés fija (y, por lo que valga, el colchón de la inversión accionaria), no le preocupa que la empresa esté bien administrada, sino que no esté tan mal administrada que no pague sus intereses, o que sea incapaz de pagar el principal cuando se venza el bono, o vuelva estos eventos más probables de lo que pensaba el tenedor del bono cuando negoció la tasa de interés. En cambio, el rendimiento del accionista está directamente relacionado con la buena administración de la empresa y con el escrúpulo con el que los administradores repartan una porción apropiada del ingreso, o sea, todo lo que esté por encima del rendimiento competitivo para los administradores por el desempeño de su función.

El peligro de la mala administración (negligencia) es menos grave que el peligro de que los administradores no traten justamente a los accionistas (deslealtad). Los administradores tienen así un fuerte incentivo para administrar bien la empresa o, si no son capaces de administrarla bien por sí mismos, para vender sus cargos a quienes sí puedan hacerlo,[34] pero es más débil su incentivo para tratar justamente a los accionistas (es decir, para maximizar el valor del acervo de la corporación por acción). Es cierto que una corporación cuyos administradores no traten justamente a los accionistas tendrá que pagar una prima si alguna vez quiere obtener capital adicional mediante una nueva emisión de acciones comunes; pero el costo del premio no será pagado (por lo menos no primordialmente) por los administradores, sino por los accionistas originales bajo la forma de una disminución de su interés.

Hay un conflicto de intereses potencial entre los administradores y los accionistas suficiente para pronosticar que los accionistas insistirían normalmente en incluir cláusulas protectoras en la escritura de la corporación.[35] El derecho de las corporaciones reduce los costos de transacción entre los accionistas y la administración al implicar en cada escritura constitutiva los derechos normales que podría esperarse que exigiera un accionista, el más importante de los cuales es el derecho de emitir votos, iguales al número de las acciones que tenga, para designar a los miembros del consejo de administración o junta directiva de la corporación. Ese consejo no administra a la empresa. Integrado de ordinario por administradores principales más (de manera creciente) administradores externos que, por trabajar a tiempo completo en otra parte, dedican sólo una atención esporádica a los asuntos de la corporación, el consejo ratifica normalmente las decisiones de los administradores. Pero selecciona al director ejecutivo de la corporación y lo remplaza si no cumple con las expectativas. La existencia del consejo evita que la administración corporativa se convierta en una oligarquía que se autoperpetúa.

Las protecciones legales de los accionistas refuerzan los instrumentos contractuales para alinear el interés propio de los administradores con el de los accionistas, tales como el de hacer depender la compensación de los administradores en gran medida del desempeño de la empresa medido por los cambios en el valor de sus acciones.[36] Discutimos antes un método más re-

[34] No se les permite que hagan esto directamente, pero hay algunos métodos indirectos de venta, especialmente como parte de una adquisición corporativa. Véase después XV.9.

[35] Hay un conflicto de intereses similar entre los beneficiarios de un fideicomiso y el administrador fiduciario (véase antes III.10), y no es sorprendente que encontremos un conjunto similar de cláusulas protectoras (obligaciones fiduciarias) en ambos contextos. Pero hay una diferencia importante. El administrador fiduciario es responsable de la diversificación de las inversiones del beneficiario. La corporación es de ordinario una empresa no diversificada; le corresponde al accionista obtener el grado de diversificación que desea mediante la compra de acciones en diversas corporaciones. Pero no adelantemos la historia. Véase después XVI.2.

[36] Véase "Symposium on Management Compensation and the Managerial Labor Market", 7 *J. Acctg. & Econ.,* 3 (1985).

buscado para el alineamiento: la adquisición apalancada. Un método más antiguo, semejante al anterior, es el pago de dividendos. Dado el tratamiento del impuesto sobre la renta para las ganancias del capital, mucho más favorable que el del ingreso ordinario, el pago de dividendos puede parecer algo extraño. Los dividendos se gravan como el ingreso ordinario; si el mismo ingreso se retiene en la corporación, el valor del conjunto de acciones de la corporación será mayor, y el accionista podrá obtener el mismo efectivo que obtendría de recibir dividendos si vende periódicamente algunas de sus acciones para obtener ganancias de capital. Con todo, el pago de dividendos ayuda a mantener en línea los intereses de los administradores con los de los accionistas, obligando a la corporación a volver con mayor frecuencia al mercado de capital para financiar nuevos proyectos, en lugar de financiarlos con las ganancias retenidas, y aumentando el riesgo de fracaso de la corporación porque la razón de deuda a capital accionario será mayor si no se retienen todas las ganancias.[37] En los países que ofrecen a los inversionistas una protección legal más débil que la ofrecida por la ley estadunidense, observamos respuestas del mercado tales como una propiedad accionaria muy concentrada antes que dispersa y primas muy elevadas por un interés accionario controlador, y respuestas legales tales como los dividendos obligatorios.[38]

Existen razones para dudar que los arreglos privados y legales para alinear el interés de los administradores con el de los accionistas, o el mercado en control corporativo, discutido a continuación, sean completamente efectivos, como veremos en XV.11 y XV.13.

XV.9. La transferencia del control corporativo

Si los administradores se están desentendiendo de los intereses de los accionistas, bajará el precio de mercado de las acciones comunes de la empresa. Los inversionistas alerta se darán cuenta de que su precio sería mayor si la empresa estuviera administrada con el objetivo de maximizar el rendimiento del accionista. Ellos podrán explotar este conocimiento en varias formas:

1. Una corporación u otro inversionista pueden tratar de persuadir a la junta directiva para que coopere en una venta o fusión que otorgará el control al adquirente. La aquiescencia de los directores y administradores de la empresa adquirida puede obtenerse mediante promesas de una compensación generosa bajo la forma de lucrativos contratos de consultoría, etc. Sin embargo, los accionistas tendrán que votar sobre la venta o fusión propuestas.

[37] Frank H. Easterbrook, "Two Agency-Cost Explanations of Dividends", 74 *Am. Econ. Rev.,* 650 (1984).

[38] Rafael La Porta *et al.,* "Law and Finance", 106 *J. Pol. Econ.,* 1113 (1998).

2. Si los administradores y el consejo de administración se resisten, un adquirente potencial podrá hacer una oferta de compra de la mayoría de las acciones vigentes de la corporación a los accionistas existentes a un precio entre el precio de mercado corriente y el precio al que aumentarían las acciones con una administración y un control apropiados. Si se acepta la oferta de compra, el adquirente tendrá votos suficientes para elegir su propio consejo de administración, el que entonces remplazará al consejo actual.
3. Un inversionista corporativo o individual puede comprar acciones suficientes: *a)* para formar una base a fin de librar una campaña por los poderes de voto de los otros accionistas y *b)* con el objeto de estar facultado para obtener excelentes ganancias del aumento en el precio de mercado de las acciones comunes de la empresa cuando se despida al consejo de administración y a los administradores anteriores.

Los jurisconsultos tradicionales del derecho corporativo hacían hincapié en la democracia corporativa (número *3*) pero no en el mercado del control corporativo. La lucha de poderes librada por un individuo que no ha adquirido una posición de propiedad sustancial se asemeja mucho al proceso democrático, pero es el método de adquisición menos viable, en parte debido a las externalidades: ¿cómo podrá tal individuo recuperar los costos de la campaña con una ganancia conmensurable al riesgo del fracaso?

La ley obstruye el uso efectivo de la oferta de compra, lo que a su vez disminuye la eficiencia de la ruta de la fusión voluntaria al reducir la presión sobre los administradores de una empresa mal manejada para que renuncien voluntariamente a su control. Por ejemplo, prohíbe que un inversionista que busque la adquisición compre una porción sustancial de las acciones de la empresa deseada sin revelar su propósito. Tal revelación tiende a aumentar el precio de la acción y, por ende, a reducir la ganancia de la adquisición y, en consecuencia, el incentivo para intentarlo; la revelación externaliza gran parte del beneficio creado por la opción de adquisición.[39]

Habría otro obstáculo si, como a veces se propone, se prohibiera que un accionista controlador, al vender sus acciones, cobrara una prima por el control de la corporación que la venta otorgaría al comprador. La teoría legal es que el accionista controlador tiene obligaciones fiduciarias para con los accionistas minoritarios. La teoría tiene mérito cuando existe un conflicto de intereses entre los accionistas mayoritarios y minoritarios. Pero, en la situación de adquisición o apoderamiento habitual, los accionistas minoritarios se verán más perjudicados que beneficiados por una regla que, al reducir el incentivo del accionista controlador para vender el control, retarda la reasig-

[39] Gregg Jarrell y Michael Bradley, "The Economic Effects of Federal and State Regulations of Cash Tender Offers", 23 *J. Law & Econ.*, 371 (1980).

nación de los activos de la corporación a personas que pueden usarlos más productivamente en beneficio de todos los accionistas.

Los blancos potenciales de las adquisiciones adoptan a menudo ciertas tácticas para volver más costosas tales adquisiciones; por ejemplo, contratos que garantizan pagos de indemnización por despido derrochadores para los administradores despedidos a resultas de una adquisición (el paracaídas dorado, como se le llama), y las píldoras de veneno. La píldora de veneno es una táctica por la que cada accionista recibe con cada acción una garantía que lo faculta para que, en el caso de que la corporación sea adquirida, venda sus acciones a la empresa adquirente al precio fijado en ella, el que de ordinario es varias veces mayor que el precio de mercado corriente de la acción. Si la oferta de compra es suficientemente generosa la píldora de veneno no causa ningún problema, porque todos los accionistas estarán felices de vender sus acciones (con las garantías anexas) a la empresa adquirente. Pero la oferta de compra óptima, desde el punto de vista del oferente, *no* atrae a todos los accionistas: eso sería demasiado caro; la meta es fijar un precio que induzca a una mínima mayoría de los accionistas a vender sus acciones. Habiendo obtenido así el control, el oferente de la compra ofrecerá entonces comprar las acciones restantes (suponiendo que no quiere ser molestado por una minoría de accionistas) pero a un precio menor porque ya no tienen otra alternativa; la diferencia del precio es la prima por el control mencionado antes. La oferta de compra en dos estratos reduce efectivamente el precio de oferta inicial, porque los accionistas competirán entre sí para estar en el primero, el estrato de precio mayor, que recibe el premio por el control. La píldora de veneno es un antídoto eficaz para las ofertas de dos estratos porque obliga a la corporación adquirente a comprar el segundo estrato a un precio mayor que el primero y así reduce las ganancias esperadas de la formulación de la oferta.

El paracaídas dorado es menos objetable que la píldora de veneno. Una garantía de pago de indemnización por despido de carácter elevado detonada por una adquisición vuelve más costosas a las adquisiciones, pero también reduce el incentivo de los administradores para resistirse a las adquisiciones; los efectos podrían ser los opuestos.

Las tácticas contra la adquisición se defienden con el argumento de que promueven la igualdad entre los accionistas, haciendo que los oferentes suavicen sus ofrecimientos, y corrigen las subvaluaciones del mercado. Estas defensas, por lo menos la primera y la tercera, son poco plausibles (especialmente la última, en vista de la tesis del mercado eficiente explorada en el capítulo siguiente). Se ha sostenido que su efecto principal es desalentar las ofertas de compra, y algunos estudios han demostrado que cuando tales oferentes son derrotados, los accionistas de la corporación deseada sufren pérdidas.[40]

[40] Véase, por ejemplo, Frank H. Easterbrook y Gregg A. Jarrell, "Do Targets Gain from Defeating Tender Offers?", 59 *N.Y.U. L. Rev.,* 277 (1984). Sin embargo, las pruebas no son conclu-

Pero las tácticas contra la adquisición requieren la aprobación de la junta directiva de la corporación: una junta que en la actualidad tiene probablemente una mayoría de directores externos. ¿Podrían estos directores estar actuando contra los intereses de la corporación? Es posible. Después de todo, tienen un conflicto de intereses, ya que si la corporación es adquirida es muy probable que pierdan sus posiciones en la junta, donde reciben excelentes honorarios. Los directores internos tienen intereses mayores aún en la continuación de la independencia de la compañía. Más desconcertante resulta que los accionistas apoyen las tácticas contra la adquisición, porque tales tácticas requieren a menudo, y reciben a menudo, el apoyo de una mayoría de los accionistas, así como de una mayoría de la junta directiva.

Si son correctos los estudios que descubren que las tácticas contra la adquisición perjudican a los accionistas, esto sugiere que no está funcionando bien el mercado del control corporativo y que su falla es grave. Aun así, el hecho de que los administradores resistan a menudo las adquisiciones que favorecerían los intereses de los accionistas no es necesariamente prueba de que haya un grave conflicto de intereses entre administradores y accionistas. Pues debemos preguntarnos cuál *debiera* ser la postura de los administradores de una empresa que es candidata a la adquisición. ¿Debieran aconsejar a los accionistas que acepten la primera oferta de adquisición que se haga, o debieran tratar de demorar la aceptación de la oferta con la esperanza de estimular una competencia entre quienes se disputan la adquisición, en la que se venderían las acciones al mejor postor? En un mercado ordinario parecería extraño que se *requiriera* a un vendedor que aceptara la primera oferta, y aunque ningún accionista está obligado a aceptar una oferta de compra, podría carecer de la información que tienen los administradores acerca de la perspectiva de una oferta mayor si rechaza la primera. Hay ciertas pruebas en el sentido de que a las empresas en la mira que resisten las ofertas de compra pero son adquiridas más tarde les va mejor que a las empresas en la mira que no resisten; aunque si resisten con éxito les va peor.[41] Comoquiera que sea, la ganancia esperada de la resistencia podría ser positiva aunque cierta porción de los resistentes lo hagan demasiado bien y no sean adquiridos en absoluto, en lugar de que sean adquiridos a un precio mayor… que es el propósito de la resistencia, si las pruebas han sido correctamente interpretadas.

Una variante de la píldora de veneno es el "aseguramiento": la corporación le promete a un postor por la corporación que, si la oferta es rechazada, la corporación lo compensará por la oportunidad perdida. Los tribunales objetan que el aseguramiento: *1)* disuade a otros adquirentes potenciales de presentar mejores ofertas, porque la corporación valdrá menos si la adquiere

yentes. Michael C. Jensen y Richard S. Ruback, "The Market for Corporate Control: The Scientific Evidence", 11 *J. Fin. Econ.,* 5, 29-40 (1983).

[41] Gregg A. Jarrell, "The Wealth Effects of Litigation by Targets: Do Interests Diverge in a Merger?", 28 *J. Law & Econ.,* 151 (1985).

quienquiera que no sea el postor que tiene el aseguramiento (debido a lo que la corporación tendrá que pagarle), y *2)* induce a la corporación a otorgar preferencia a un postor que preservará a los administradores existentes aunque los activos de la corporación valgan efectivamente más en las manos de otro adquirente. Pero estas objeciones no son decisivas necesariamente.[42] Un postor a quien se le prometiera una compensación si su postura fuese rechazada ofrecerá un precio mayor, y si su postura es aceptada la corporación no tendrá que pagar ninguna compensación. De este modo, aun las juntas directivas leales —las que estén decididas a maximizar los intereses de los accionistas antes que los suyos propios— podrían utilizar los aseguramientos. Una junta desleal a la que se le prohíba usar los aseguramientos podría alcanzar el mismo objetivo por otros medios; por ejemplo, negándose a aceptar las posturas del mejor postor sin pagos laterales a los administradores.

La mayoría de las adquisiciones son amistosas; pero hay pruebas de que la empresa adquirente (el postor triunfante) no obtiene con mucha frecuencia una ganancia con la adquisición. Ésta podría ser otra manifestación de los costos de agencia, específicamente un reflejo de los incentivos de los administradores para la diversificación y el crecimiento.[43] Los administradores, en virtud de que tienen un capital específico de empresa (mientras que los accionistas carecen de él), poseen un mayor incentivo para diversificarse con la finalidad de reducir el riesgo de la quiebra y, probablemente, la pérdida consecuente de su empleo, y, asimismo, un mayor incentivo para incrementar el tamaño de la empresa, tanto para reducir el riesgo de colapso de la firma como porque resulta más fácil "ocultar" las rentas gerenciales entre más grande y más compleja sea la empresa y más difícil sea, por ende, monitorearla.

XV.10. Expulsiones corporativas y la competencia en la laxitud

En los términos de la ley de corporaciones de Delaware, la mayoría puede obligar a la minoría a vender sus acciones a la corporación sin tener que demostrar ningún propósito comercial para la acción. Los accionistas minoritarios que estén descontentos con el precio ofrecido por la mayoría para sus acciones pueden demandar una evaluación judicial del valor de tales acciones, de modo que la mayoría tiene en efecto un poder de dominio eminente respecto de las acciones de la minoría. El que deba permitirse esta transferencia forzada depende en parte de que haya cualquier base razonable para que una mayoría desee eliminar a la minoría. La hay. Los activos de una corporación podrían ser más valiosos si las acciones de la corporación se

[42] Stephen Fraidin y Jon D. Hanson, "Toward Unlocking Lockups", 103 *Yale L. J.,* 1739 (1994).

[43] Véase Andrei Shleifer y Robert W. Vishny, "A Survey of Corporate Governance", 52 *J. Fin.,* 737, 747-748 (1997). Más datos sobre estos incentivos se hallan abajo en XV.11.

mantienen privadamente en lugar de que estén registradas en los términos de las leyes de valores y se negocien en una de las bolsas organizadas, y, por lo tanto, estén sujetas a extensas regulaciones gubernamentales y de la bolsa de valores. El cumplimiento de estas regulaciones es costoso y una fuente de riesgos legales, pero además dificulta que una empresa opere en secreto, lo que podría impedir que la empresa se aproveche de ciertas oportunidades de negocios. Por otra parte, el hecho de operar en privado genera una unión más estrecha de la propiedad y el control, lo que reduce los costos de agencia. Por lo tanto, si un accionista mayoritario (o varios accionistas mayoritarios) decide que su corporación pública sería más valiosa si se volviera privada, deberá persuadir a la mayoría de los demás accionistas para que renuncien a sus acciones. Ésta parecería una negociación fácil porque, por hipótesis, la corporación valdría más privada que pública. Pero, sobre todo cuando los accionistas minoritarios son numerosos, hay un grave problema de renuencia de una clase que ya debiera ser muy familiar para el lector.

La preocupación por las expulsiones se relaciona con la noción prevaleciente de que los estados compiten para atraer corporaciones promulgando leyes de corporaciones que proveen protecciones inadecuadas para los acreedores y para algunos de los accionistas o para la totalidad de ellos, y de que Delaware —el estado de incorporación de casi la mitad de las compañías listadas en la Bolsa de Valores de Nueva York— ha ganado la competencia por la laxitud de su ley de corporaciones. Hay propuestas perennes para que se elimine esta competencia requiriendo que las corporaciones más grandes se constituyan de acuerdo con una ley federal uniforme de las corporaciones. Pero los propios accionistas minoritarios podrían verse perjudicados por una disposición que, al otorgarles el poder de bloquear las expulsiones, pudiera impedir que las corporaciones obtuvieran mayores ganancias volviéndose privadas.[44] Los propios accionistas minoritarios se verían perjudicados ya que su compensación por la entrega de sus acciones sería mayor entre mayores fueran los beneficios para la corporación de convertirse en privada.

[44] Un estudio empírico ha descubierto que de hecho las expulsiones no explotan a los accionistas minoritarios. Harry DeAngelo, Linda DeAngelo y Edward M. Rice, "Going Private: Minority Freezeouts and Stockholder Wealth", 27 *J. Law & Econ.,* 367 (1984).

¿Debería el valor de la empresa como un negocio en marcha con buen funcionamiento, tal como se le evalúe, incluir el valor presente del aumento anticipado en las ganancias por volverse privada? ¿Debe la corporación rembolsar los gastos de la litigación de la minoría en el proceso de evaluación? Sí y sí. Si al elegir la privatización pudiera la mayoría expropiar la parte proporcional de la minoría de alguna nueva oportunidad corporativa, se crearía un incentivo para volver privada a la empresa que superaría los beneficios en materia de eficiencia derivados de esa transmutación (esto es, la mayoría contaría tanto con razones relativas a la búsqueda de la renta como con razones que competen a la asignación). De modo similar, si los gastos de la minoría no son rembolsados, la mayoría podría tener la facultad de forzarlos a aceptar un precio por debajo del poder real de sus participaciones, lo que tendría los mismos efectos de incentivo que lo anterior.

De acuerdo con el modelo más simple de comportamiento corporativo, las corporaciones no se incorporarían a un estado que no proveyera ninguna protección para los acreedores o los accionistas porque, si lo hicieran, tendrían que pagar a los acreedores tasas de interés muy elevadas o bien tendrían que aceptar, en sus convenios de préstamos, cláusulas protectoras muy elaboradas más allá de las provistas por la ley del estado. Y una corporación privada que deseara volverse pública pero ofreciera sólo una minoría de sus acciones vigentes al público inversionista no desearía estar gobernada por una cláusula de expulsión injusta para los accionistas minoritarios, pues los accionistas potenciales advertirían que, como accionistas minoritarios en una corporación que tiene una mayoría compacta, serían vulnerables a la expulsión. Sería difícil compensarlos anticipadamente por afrontar ese riesgo porque el riesgo se halla bajo el control de la mayoría; no es un evento externo. Existiría el mismo tipo de problema de azar o riesgo moral que el surgido si las personas pudieran comprar un seguro contra incendios que ellas hubieran causado deliberadamente.

¿Significa esto que Delaware ha ganado la batalla por las incorporaciones debido a que una porción tan grande de sus ingresos procede de cuotas corporativas estatutarias que las corporaciones saben que tiene un interés financiero en mantener vigente un cuerpo considerable de jurisprudencia interpretativa cuya aplicación corre a cargo de jueces especializados de indudable competencia?[45] Quizá no. Hay pruebas de que muchos estados no compiten por incorporaciones realizadas por corporaciones con sede en otros estados y de que los que sí compiten lo hacen ofreciendo a éstas una férrea protección contra las adquisiciones o apoderamientos de signo hostil, la cual bien puede (si bien esto es incierto: véase XV.9 antes) ser dañina para los accionistas.[46]

XV.11. El problema de la compensación

En el modelo económico más simple de compensación, a un trabajador o empleado, hasta el nivel del funcionario ejecutivo de primer nivel en una corporación gigantesca, se le paga su producto marginal: esencialmente, su aportación al ingreso neto de la compañía. Aun así, uno observa que los salarios varían entre empleados del mismo nivel en la misma compañía por mucho menos que la diferencia entre sus contribuciones a la rentabilidad de la compañía, y que los empleados que realizan un trabajo satisfactorio pueden esperar incrementos anuales reales (esto es, ajustados a la inflación) en sus salarios a lo largo de sus carreras dentro de la empresa, pese a que sus con-

[45] Véase Roberta Romano, "Competition for Corporate Charters and the Lesson of Takeovers Statutes", 61 *Fordham L. Rev.,* 843 (1993).

[46] Véase Lucian Arye Bebchuk y Alma Cohen, "Firms' Decisions Where to Incorporate", 46 *J. Law & Econ.,* 383 (2003).

tribuciones no estarán aumentando a la misma velocidad y, a la larga, no aumentarán de ninguna manera. El primer fenómeno puede reflejar la dificultad de medir las contribuciones individuales a la producción de equipo. A pesar de la dificultad, a menudo es patente (tanto para los demás miembros del equipo como para la administración) que un miembro del equipo es el mejor; por consiguiente, éste recibe una promoción y un salario más elevado, y de esta manera se recompensan al menos las diferencias marcadas en la contribución. El segundo fenómeno puede ser concebido para resolver el problema del último periodo (véase antes XII.10).

No obstante, ¿puede el modelo estándar de la compensación, incluso cuando se le refina hasta este punto, explicar la compensación de los directores corporativos? Los directores ejecutivos o generales en los Estados Unidos reciben casi el doble de pago en promedio que el de sus contrapartes en otros países. No obstante, esto no se debe a que los ciudadanos estadunidenses ganen en todos los niveles más que los de otros países; la diferencia entre los salarios en los Estados Unidos y los del extranjero es mucho más baja si nos situamos en el nivel inferior al de los directores ejecutivos. La causa aproximada de la diferencia que hay en la forma de compensar a estos últimos es que los salarios constituyen una fracción mucho más pequeña en la retribución (menos de la mitad) de los directores ejecutivos estadunidenses que en el caso de los salarios recibidos por sus correspondientes extranjeros: el resto consiste en bonos, pero principalmente en opciones en materia de valores o acciones. Las opciones relativas a valores y otros métodos de compensación basados en incentivos dotan de (una variación de) riesgo a los ingresos de los directores ejecutivos. Así se refuerza el riesgo inherente al hecho de que gran parte del capital humano de un director general puede ser específico de su firma. En virtud de que los directores ejecutivos, diferenciándose de los empresarios, no son afectos al riesgo, han de exigir un salario más elevado si éste conlleva un componente sustancialmente riesgoso. Ello puede explicar *en parte* la diferencia entre la compensación recibida por los directores ejecutivos estadunidenses y los extranjeros, pero no su totalidad, ni siquiera una gran proporción de ella, en especial debido a que la rotación de empleo al nivel en que se encuentran ellos es de hecho más grande en Europa que en los Estados Unidos.

Por supuesto, si los accionistas desean que el director ejecutivo realmente confronte enormes riesgos, tendrán que pagarle una compensación de veras enorme para lograr que su retribución esperada sea equivalente a la de sus pares. Algunas compañías de capital privado adoptan tal estrategia, pero esto no puede explicar la diferencia internacional en la compensación promedio.

Una explicación más convincente es que la propiedad de valores está más concentrada en el extranjero que en los Estados Unidos. Los accionistas poseen un mayor incentivo para monitorear el desempeño de los administradores de su compañía entre más considerable sea lo que tienen en juego. Y entre más efectivo sea el monitoreo de los accionistas, menor será la necesidad de

una compensación basada en incentivos: la zanahoria es sustituida por el garrote. Así, tal vez los ingresos de los directores ejecutivos estadunidenses resultan excesivos en el sentido de que son más altos de lo que serían en un mercado efectivamente competitivo para los gerentes corporativos.[47] Esta inferencia se respalda por el hecho de que en sus salarios el componente "incentivo" de sus ingresos que es más significativo —las opciones sobre valores o acciones— no se correlaciona adecuadamente con la aportación del director ejecutivo al valor de su compañía. Son muchas las cosas que mueven a las reservas de una compañía aparte de las decisiones de su director general. Ligar su ingreso al valor de las acciones de su compañía es algo así como ligar el salario del presidente de los Estados Unidos al producto nacional bruto. La analogía resulta cercana en el caso de una industria como la petrolera. Las ganancias de una compañía petrolera son en una gran medida una función del precio del petróleo, sobre el que las compañías tienen poco control. Por añadidura, ligar la compensación de un ejecutivo a las reservas de la corporación genera un incentivo para manipular el precio de las acciones, y este incentivo ha sido responsable de una serie de debacles en el campo financiero.

La elección de las opciones en materia de valores como el método principal para otorgar una compensación no salarial a los directores ejecutivos puede guardar relación con el hecho de que, hasta fecha muy reciente, entre las reglas de contabilidad no se incluía el requisito de que el ingreso generado por estas opciones, a diferencia del ingreso del salario o por bonos, se reportara como un gasto corporativo; y usualmente no se daba cuenta de él. Por supuesto, los analistas de los activos y los accionistas poseedores de una cuota lo suficientemente cuantiosa para permitirles seguir de cerca los asuntos de las compañías en las que han invertido pueden calcular el gasto representado por las opciones sobre valores, pero el público ordinario no puede hacerlo; y en el siguiente capítulo veremos pruebas de que los mercados de capital no son tan competitivos como los economistas solían pensar que lo son.

Uno pudiera pensar que la competencia en los mercados de productos de la corporación pondría un coto a la avaricia gerencial porque ésta incrementa los costos de la corporación. Pero el problema de los costos de agencia es inherente a la estructura de cualquier empresa de grandes proporciones y, por ello, es probable que acose a todos los competidores de envergadura en un mercado y, por lo tanto, no sea eliminado por la competencia aun cuando los mercados en los que las empresas venden sus productos o servicios sean competitivos en la mayor medida. Asimismo, el efecto que tiene la compensación excesiva a los directores ejecutivos en el bienestar de las empresas es indirecto. El efecto directo consiste simplemente en transferir riqueza de los accionistas a los administradores, si bien el efecto indirecto (más allá de cua-

[47] Como se sostiene con argumentos sólidos en Lucian A. Bebchuk y Jesse M. Fried, *Pay without Performance* (2004).

lesquiera distorsiones habidas en el comportamiento gerencial) es el de incrementar el costo de obtener nuevo capital accionario para la empresa.

Otro control imperfecto del exceso en la compensación corporativa es el de la opinión pública. A las compañías no les costea relegar completamente a la opinión pública, dado que ésta, de ser adversa, puede empoderar alguna medida legislativa o regulatoria que resulte dañina para ellas o para una industria. Las compañías hacen uso de ingeniosas estratagemas mediante las cuales una compensación al director ejecutivo que a una persona común y corriente le daría la impresión de ser tremendamente excesiva se le oculta al público; al exhibir a la luz del día tales estratagemas y, conjuntamente, al divulgar de costa a costa los escándalos corporativos, se puede ejercer una presión a la baja sobre la compensación provista a dichos directores.

El costo que significa para una compañía de considerable magnitud incluso otorgar una paga excesiva a un director general es tan reducido una vez que se divide entre los accionistas que no habrá uno solo de éstos (suponiendo que se trata de propietarios por separado) que tenga ningún incentivo para hacer cualquier cosa al respecto. El incentivo que tenga la junta de directivos será igualmente débil, en parte porque probablemente la junta está dominada por ejecutivos de negocios muy bien remunerados, incluidos algunos directores ejecutivos o generales, que poseen un interés económico en mantener salarios elevados en la esfera corporativa y presentan una tendencia psicológica natural a creer que esos salarios reflejan fielmente el valor intrínseco de sus receptores. Asimismo, una junta que contratara al segundo mejor candidato por la razón de que le costaría a la compañía mucho menos quedaría sujeta a críticas severas si el nuevo director ejecutivo resultara ser un fiasco. La junta se protege a sí misma pagando muchos dólares a fin de conseguir al mejor entre los mejores; entonces, si este individuo da traspiés colosales, será más difícil que la culpa recaiga en la junta.

Las empresas consultivas que las juntas directivas contratan para que las aconsejen sobre cómo seleccionar y compensar a un director ejecutivo se pliegan mansamente a los instintos autoprotectores de los directivos recomendando de forma característica que la junta contrate como director general a alguien que habrá de demandar una compensación en por lo menos el septuagésimo quinto percentil de los directores generales de las compañías que, según ha identificado el consultor, son comparables a la empresa de la junta. A resultas de ello, la compensación se dispara hasta los cielos; el septuagésimo quinto percentil de antaño deviene el quincuagésimo percentil. El conflicto de intereses en que se inmiscuyen los consultores a propósito del tema en cuestión se ve agravado por el hecho de que (tal como ocurre con las empresas de contabilidad) ellos le venden a la empresa otros servicios de consultoría.

En contraposición a la teoría de la compensación ejecutiva basada en la búsqueda de renta, se ha sostenido que el incremento habido a lo largo del tiempo en la retribución dada a los directores ejecutivos es una función del cre-

cimiento del valor de mercado que experimentan las compañías.[48] Aquella teoría, que guarda consistencia con la teoría de la compensación fundada en el producto marginal elemental, predica que el principal ejecutivo de una compañía más valuada es más productivo, puesto que, si él aumenta el valor de la empresa por un determinado porcentaje, el incremento absoluto en el valor de la firma será más elevado entre más valuada resulte ésta. (Si existen dos administradores parejamente habilidosos y uno administra una tienda de abarrotes mientras que el otro administra la IBM, el que está generando un valor más grande es probablemente el segundo.) Con todo, otra explicación es que, entre mayor sea el valor de mercado de la compañía, más fácil será "ocultar" la compensación dada a los ejecutivos del máximo nivel. Si un aumento de 10% en el valor de la empresa se asocia con un aumento de 3% en la compensación otorgada a su principal ejecutivo, entonces el porcentaje del valor de la empresa que va a parar a manos de este individuo habrá disminuido. Ésta puede ser otra razón por la que muchas fusiones no aciertan a incrementar las ganancias por acción. El valor agregado de la empresa será mayor, lo que permite que el incremento en la compensación del director ejecutivo pueda ser ocultado con más facilidad.

Otros estudiosos del tema de la compensación argumentan que, desde el punto de vista del accionista, la habilidad "gerencial" que resulta crítica no es la habilidad para administrar una gran empresa de manera eficiente sino, antes bien, la habilidad, acaso más escasa, de hacer uso de las relaciones personales y los tejemanejes contables con la finalidad de crear una burbuja en las reservas de la corporación.[49] Entre más valor adquieran las reservas de la corporación, menor será el costo para ésta de adquirir capital adicional —ya sea mediante el apoderamiento o adquisición de otras empresas o a través de la emisión de nuevas acciones— y, similarmente, el costo de atraer a buenos ejecutivos y empleados de otra índole. Si este análisis es correcto, resulta eficiente fundamentar la compensación de un director general en el desempeño observado por las reservas de la corporación en vez de en los "principios básicos" de la compañía, tales como sus ganancias, ingresos y costos.

No puede darse por sentado que la competencia en los mercados de productos y de capital de las corporaciones elimine la ineptitud y el mal comportamiento a nivel gerencial. El problema de los costos de agencia es inherente a la estructura de cualquier empresa de magnitud y, por lo tanto, probablemente azota a todos y cada uno de los competidores de envergadura en cada mercado, por lo que no se podrá eliminar con la competencia.

[48] Xavier Gabaix y Augustin Landier, "Why Has CEO Pay Increased So Much?", 123 *Q. J. Econ.*, 49 (2008).

[49] Andrei Shlejfer, "Does Competition Destroy Ethical Behavior?", 94 *Am. Econ. Rev. Papers & Proceedings*, 414 (mayo de 2004).

XV.12. La negociación de las personas que poseen información confidencialde la empresa y el problema de la recompensa empresarial

Los administradores y otros empleados dentro de la empresa que poseen información confidencial tienen prohibido utilizar información sustancial aún no revelada a otros accionistas o al mundo exterior a fin de obtener ganancias negociando con las acciones de la empresa. El comercio con la información privilegiada ha atraído poderosamente la atención en los últimos tiempos a raíz de que en la ciudad de Nueva York los fiscales realizaran esfuerzos sin precedentes para combatirlo con todo vigor, en una campaña que ha revelado que este fenómeno está extendido en grado extremo y es muy lucrativo; así es como se confirma la impresión cada vez mayor de que la corrupción está expandida por toda la industria financiera, por la razón más bien obvia de que es en elevada medida rentable y, por consiguiente, atrae a una cantidad desproporcionada de gente en extremo codiciosa.

Aun así, esta prohibición ha sido criticada alegando que la negociación de los que están dentro de la empresa y poseen información privilegiada es una táctica de incentivo potencialmente importante.[50] Sobre todo si, como es plausible, los administradores sienten por el riesgo una aversión mayor que la de los accionistas, el hecho de que se permita la circulación de información confidencial, al alentar a los administradores a asumir riesgos mayores a nombre de la corporación, podría alinear los incentivos de éstos más estrechamente con los incentivos de los accionistas. También podría permitir que la corporación pagara sueldos menores a los administradores. Y podría conducir a determinar de manera más correcta el precio de las acciones de una corporación, al permitir que la información que poseen algunas personas que están dentro de ella, pero nadie más, afectara la negociación de las acciones y, por ende, su precio; en esa forma, los datos quedarían efectivamente disponibles para otros inversionistas. Sin embargo, la información privilegiada no remunera a la administración eficiente como tal. Remunera la posesión de información confidencial, ya sea favorable o desfavorable para las perspectivas de la corporación. Y coloca a los inversionistas que no tienen (o se rehúsan, movidos ora por el miedo o por sus principios, a entablar negociaciones al respecto) acceso a la información confidencial en desventaja, y en esta medida desalienta el comercio y la investigación sobre inversiones y reduce la liquidez.

[50] La exposición clásica de esta concepción es Henry G. Manne, *Insider Trading and the Stock Market* (1966).

XV.13. Los escándalos corporativos

En años recientes, una cantidad sorprendente de administradores de alto nivel pertenecientes a corporaciones de gran magnitud han sido procesados debido a fraudes corporativos que en algunos casos (como en los de Enron y WorldCom) llevaron a inmensas pérdidas de los accionistas. La mala conducta de estos administradores, que ha sido muy difundida por los medios, sirvió de acicate para promulgar la Ley Sarbanes-Oxley de 2002,[51] que aumentó los castigos por fraude corporativo, volvió más rígidos los estándares para la auditoría (con lo que prohibió a las empresas comprar servicios de consultoría por parte de sus auditores), magnificó la autoridad de los directores independientes, limitó la práctica sostenida por las corporaciones de prestarles dinero a sus ejecutivos y, por estos y otros medios, persiguió la meta de hacer que los administradores corporativos sirvieran más fielmente como agentes de sus principales (nominales), los accionistas. La preocupación por el mal comportamiento de los administradores corporativos se avivó con el colapso financiero de 2008 y la crisis económica mundial que le siguió; pero ése es un tema del siguiente capítulo.

Varios de los tipos recurrentes de mala conducta que se han denunciado se relacionan directamente con el asunto de la compensación de los directores ejecutivos. Son:

1. La incapacidad de considerar el valor de las opciones en materia de reservas como una responsabilidad corporativa, es decir, de gastarlas, ocultando así de la vista de los accionistas la compensación realmente recibida por los administradores.
2. La incapacidad de revelar en las declaraciones financieras de la compañía el costo de los beneficios no pecuniarios para los funcionarios, incluidos los funcionarios retirados, tales como el uso de aviones y residencias de la compañía, lo que produce el mismo efecto que el señalado en el punto anterior.
3. El no dar a conocer los préstamos que las corporaciones hacen a sus funcionarios, o el subsiguiente perdón concedido a los préstamos, nuevamente con el mismo efecto.
4. La supervisión inadecuada de la administración por la junta directiva, particularmente sus comités dedicados a auditoría y compensación.
5. La emisión obsequiosa de las opciones sobre valores en bien de los ejecutivos corporativos, una práctica que, según se señala, ha llevado a que la administración se preocupe en forma indebida por las fluc-

[51] Pub. L. No. 1 07-204, 116 Stat. 745 (codificado en secciones dispersas de los títulos 15 y 18 del Código de los Estados Unidos [U.S. Code]).

tuaciones a corto plazo en los precios de las acciones y anhele manipularlas.

Más aún, algunos de los otros abusos corporativos recurrentes están vinculados, así sea indirectamente, a la compensación de los ejecutivos: el de inflar el valor de los activos de la corporación al transferir algunos de ellos a precios acrecentados a las así llamadas "entidades de propósitos especiales", así como los conflictos de intereses surgidos cuando los ejecutivos corporativos son remunerados por administrar las entidades controladas, y la incapacidad de las empresas de contabilidad para auditar a sus clientes con cuidado, lo que quizá se origina por el conflicto de intereses inherente en el hecho de que los clientes pagan la cuota de auditoría y al hecho de que también pueden estar pagándole a su auditor por servicios de consultoría, con lo que se incrementa la dependencia del auditor respecto de la buena opinión que tiene de él la compañía.

El elemento común en la lista es el enriquecimiento de los funcionarios corporativos debido a que se les conceden beneficios corporativos que no se dejan a la vista o se les revelan insuficientemente a los dueños nominales de la corporación, los accionistas, o a que se les ocultan a estos últimos malas noticias que provocarían el desplome del precio de las acciones de la corporación y, con ello, el derrumbe de la compensación recibida por los funcionarios. En suma, la acusación es que se ha registrado un saqueo corporativo a manos de la administración; de esta manera, el asunto del escándalo se relaciona estrechamente con el asunto de la compensación excesiva. La primera pregunta que un economista debiera lanzar es si semejante "saqueo" impone costos sociales o meramente redistribuye la riqueza de los accionistas hacia los funcionarios. Hace ambas cosas. La estrategia proyectada convierte al menos una parte de las ganancias esperadas en costos, dado que los funcionarios gastan recursos —el equivalente al costo de las herramientas empleadas por los ladrones— a fin de atraer para sí lo saqueado. Y las víctimas potenciales tienen un incentivo para incurrir en costos con el propósito de prevenirse contra la perspectiva de convertirse en las víctimas; de este modo, los inversionistas pueden ser desviados de las reservas corporativas a tipos de inversión que suministran un rendimiento social de menor escala.

El mal comportamiento corporativo no es ninguna novedad, pero se acrecentó a partir de mediados de la década de 1990, un periodo caracterizado por un dramático auge en el mercado de valores seguido por una quiebra igualmente dramática. Cuando los valores de una corporación se elevan con rapidez, la preocupación de los accionistas y los directores por el desvío de ganancias rumbo a los funcionarios se atenúa. Es probable que el desvío no tenga consecuencias (aun cuando se eleve en términos absolutos, puede estar cayendo en términos porcentuales si el mercado de valores va al alza), y los funcionarios pueden afirmar con mayor o menos plausibilidad que el alza se ha de atribuir, en parte al menos, a sus propios esfuerzos, y de esta manera

se justifica por razones de incentivos otorgarles una mayor recompensa. Cuando tiene lugar el desplome o quiebra, los funcionarios cuentan con un incentivo para emplear tácticas que pospondrán el colapso en lo referente al valor de las reservas de la compañía y, así, amortiguarán el impacto de ese colapso en sus riquezas personales.

Los instrumentos mediante los cuales el derecho corporativo se empeña en reducir la divergencia entre los intereses de los accionistas y los de los administradores —requerir que exista una junta con directivos elegidos por los accionistas a fin de supervisar a los funcionarios, que la corporación contrate a auditores independientes para corroborar que sus declaraciones financieras se apeguen a la verdad y que sus funcionarios asuman de cara a los accionistas obligaciones fiduciarias con cumplimiento legalmente exigible—, más el hecho de que los analistas de valores empleados por las casas de corredores y los bancos de inversiones monitorean el desempeño de las corporaciones en las que invierten sus clientes, como lo hacen los bancos y otras instituciones que otorgan créditos, dependen por igual para su eficacia de información que está en una gran proporción bajo el control de los funcionarios corporativos. Los directivos son vulnerables porque dedican sólo una pequeña fracción de su tiempo a la compañía. Y, si son directivos ubicados dentro de la compañía (esto es, si son empleados de tiempo completo de ésta), enfrentan un conflicto de intereses, mientras que si son externos (independientes) tienen menos acceso a la información sobre la compañía, lo que sin duda es una razón de que no haya una prueba concluyente de que el rendimiento corporativo es una función del porcentaje de directivos independientes incluidos en la junta. Los administradores corporativos influyen sobre la selección tanto de los directivos externos como de los internos; contratan y remuneran a los auditores y blanden contratos de consultoría enfrente de los auditores que también ofrecen servicios de consultoría, y pueden influir en los reportes de analistas al enviar cuotas financieras a bancos de inversión cuyos redactores les brindan a sus compañías una calificación resplandeciente.

Ahora bien, ¿qué tan sólida es la argumentación a favor de incrementar los controles regulatorios a la manera de los que impone la Ley Sarbanes-Oxley?[52] Por ejemplo, la inflación efectuada en los activos corporativos al transferirlos a precios hinchados a entidades de propósitos especiales (uno de los principales abusos de Enron) constituye una forma de fraude cuando es ocultada de la vista de los inversionistas de la corporación cesionista. Empero, el fraude ha sido desde hace largo tiempo un delito, y la persecución exitosa de los ejecutivos de Enron sugiere que desde antes del surgimiento de esta ley había instrumentos legales oportunos para vérselas con semejante

[52] Como se argumenta en Larry E. Ribstein, "Market *vs.* Regulatory Responses to Corporate Fraud: A Critique of the Sarbanes-Oxley Act of 2002", 28 *J. Corp. L.*, 1 (2002); Roberta Romano, "The Sarbanes-Oxley Act and the Making of Quack Corporate Governance", 114 *Yale L. J.*, 1521 (2005).

conducta. Sería un error prohibir todas las transferencias de activos a entidades de propósitos especiales porque éstas proporcionan beneficios económicos genuinos, a saber, la desintermediación de la deuda (es decir, la eliminación de los eslabones de intermediarios entre el último prestamista y el último prestatario). Y, muy aparte de la Ley Sarbanes-Oxley, sería posible, invocando lo que en derecho fiscal se conoce como la doctrina de "la sustancia sobre la forma", proscribir tales transferencias cuando en realidad no tienen ningún propósito económico, sino que están sencillamente diseñadas para pasar la deuda corporativa de un bolsillo a otro con la esperanza de que a los inversionistas les sea difícil localizar al segundo bolsillo.

En lo que respecta a la recepción por compañías de contabilidad de cuotas por servicios de consultoría, así como por auditorías, el auditor de Enron, Arthur Andersen, había realizado esa separación antes del escándalo de Enron. No hay ninguna razón contundente para exigir dicha separación legalmente. Debería bastar con requerir que la corporación revele a los inversionistas los términos de las relaciones que establece con sus auditores, dejando a los inversionistas la opción de penalizar a una corporación disminuyendo el precio de sus reservas si creen que el auditor ha sido "comprado". Lo que es más importante (dadas las dudas sobre la eficiencia que alberga la respuesta del mercado de valores a tal información), resulta difícil ver quién aparte de las firmas auditadas le pagaría al auditor, o cómo podría impedirse que los clientes pudieran reciclar las cuotas que ahora les pagan a los auditores por servicios de consultoría como cuotas más elevadas por auditoría; de manera que subsiste el conflicto de intereses fundamental.

Las opciones en materia de valores no debieran prohibirse, ya que algunas observan una tendencia a alinear la compensación a los administradores con el desempeño de la empresa.[53] Su eficacia y su carácter apropiado, al igual que su magnitud y su impacto neto sobre el valor de la empresa, recibirán en adelante, a raíz de los escándalos difundidos a diestra y siniestra, un mayor escrutinio del mercado… como debiera ocurrir, pues son un instrumento muy torpe para incentivar a los administradores. Tendría más sentido basar las opciones sobre valores en el desempeño de las reservas de una compañía en relación no con un periodo base, sino con las reservas de las otras compañías en la misma industria.

Resulta difícil de creer que gran parte del asunto depende de si las opciones sobre valores son contempladas, de la forma en que debieran serlo, como un gasto corporativo. La información concerniente al número de opciones en materia de valores concedidas y a quién se le concedieron y cuál es el precio de ejercicio es pública; ahora, al menos, los analistas se han sensibilizado acerca de la cuestión.

[53] Brian J. Hall y Jeffrey B. Liebman, "Are CEOs Really Paid Like Bureaucrats", 113 *Q. J. Econ.,* 653 (1998).

No es mucho lo que puede hacerse para mejorar la supervisión del cuerpo gerencial a cargo de la junta directiva. Los directivos carecen de la información y los incentivos que necesitarían para mantener a raya a los administradores. Su función principal, además de seleccionar a los directores ejecutivos, consiste en intervenir en situaciones de crisis (por lo común remplazando a éstos).

XV.14. La discreción gerencial y las responsabilidades sociales de la corporación

Si no hubiera competencia en los mercados de productos, ningún mercado para el control corporativo, ninguna gobernanza a manos de los directores y accionistas y ninguna ley de las obligaciones fiduciarias,[54] los administradores corporativos no se verían obligados a maximizar las ganancias de la corporación. Ellos maximizarían su propia utilidad, presumiblemente quedándose con todo el ingreso neto de la corporación bajo la forma de sueldos, o quizá liquidando la corporación y repartiéndose sus activos entre ellos mismos como bonos. Nadie piensa que los administradores corporativos tengan tanta libertad que puedan comportarse de este modo. Pero algunos economistas están convencidos de que en las corporaciones públicas grandes los administradores están tan poco limitados que al menos pueden sustituir ese escenario con maximandos tales como las ventas, o el crecimiento, o aun el poder personal que pueden aproximarlos a la maximización de la utilidad personal en mayor medida que la maximización de las ganancias.[55] ¿Pero son tales

[54] Los administradores son considerados fiduciarios de sus accionistas; su obligación fiduciaria puede hacerse cumplir mediante juicios derivados de los accionistas, que son juicios promovidos en nombre y a cuenta de la corporación, pero por un accionista antes que por la corporación misma (es decir, por los administradores). Por lo que toca a la economía de los juicios derivados, véanse Daniel R. Fischel y Michael Bradley, "The Role of Liability Rules and the Derivative Suit in Corporate Law: A Theoretical and Empirical Analysis", 71 *Cornell L. Rev.,* 261 (1986); Roberta Romano, "The Shareholder Suit: Litigation without Foundation?", 7 *J. Law, Econ. & Organization,* 55 (1991); Reinier Kraakman, Hyun Park y Steven Shavell, "When Are Shareholders Suits in Shareholder Interests?", 82 *Geo. L. J.,* 1733 (1994).

[55] William J. Baumol, *Business Behavior, Value and Growth* (ed. rev., 1967); Oliver E. Williamson, *The Economics of Discretionary Behavior: Managerial Objectives in a Theory of the Firm* (1967). Algunos estudios han descubierto, en respaldo de esta posición, que las empresas dominadas o controladas por un solo accionista son consistentemente más rentables que las empresas de propiedad dispersa (controladas por los administradores), aunque sólo en los casos donde es débil la competencia en los mercados de productos de la empresa. Otros estudios no han descubierto ningún efecto del control del propietario sobre el desempeño de la empresa. Véanse algunos estudios ilustrativos de ambos lados de la cuestión en James L. Bothwell, "Profitability, Risk, and the Separation of Ownership from Control", 28 *J. Ind Econ.,* 303 (1980); B. J. Campsey y R. F. DeMong, "The Influence of Control on Financial Management: Further Evidence", 18 *Rev. Bus. & Econ. Research,* 60 (1983); Robert L. Conn, "Merger Pricing Policies by Owner-Controlled *versus* Manager-Controlled Firms", 28 *J. Indus. Econ.*, 427 (1980); Cynthia A. Glassman y Stephen A. Rhoades, "Owner *vs.* Manager Control Effects on Bank Performance", 62 *Rev. Econ. & Stat.,*

maximandos realmente distintos de las ganancias? Supongamos que el crecimiento de las ventas corporativas es lo que los administradores desean maximizar, porque mientras más grande es la empresa más grande puede ser su compensación sin atraer la atención hostil de los accionistas. Para crecer rápidamente, una empresa debe generar grandes cantidades de efectivo con sus operaciones corrientes u obtener dinero de los mercados de capital. Para ambos propósitos, necesita grandes ganancias (corrientes o esperadas). La maximización del crecimiento se mezcla así inconscientemente con la maximización de las ganancias. Consideremos ahora el poder personal como un maximando. El ejecutivo corporativo más poderoso es aquel que controla una empresa muy rentable. Es el que tiene menores probabilidades de encontrarse con la crítica de los accionistas, ya no digamos con una amenaza de adquisición. Además, las grandes ganancias generan capital (y permiten que se obtenga capital adicional en términos favorables) que el ejecutivo puede usar para otros proyectos.

Mientras que algunas personas critican a la corporación moderna por no tratar de maximizar las ganancias con suficiente tesón, otras la critican por convertir la maximización de las ganancias en su única meta. Desean que las corporaciones dediquen una parte de sus ingresos a necesidades sociales tales como el control de la contaminación o la capacitación de miembros de minorías en desventaja.[56] Con todo, la filantropía puede tener efectos contraproducentes; es incluso posible que la reducción corporativa "voluntaria" de la contaminación, coordinada a través de asociaciones mercantiles a fin de minimizar el gorroneo, dado que una empresa que no está en consonancia con la campaña voluntaria tendrá costos más bajos y por ello una ventaja competitiva, podría en realidad *incrementar* la cantidad de contaminación. La demanda de regulación de la contaminación depende, en parte, de la cantidad de contaminación. Si esa cantidad se reduce mediante la acción voluntaria de los contaminadores, es posible que la demanda de regulación baje más allá del punto crítico a partir del cual los proponentes pueden superar el problema del gorroneo que afecta a la organización para la acción política.[57]

No obstante, el punto básico es que en los mercados competitivos un compromiso sostenido con cualquier meta distinta de la rentabilidad hará que la empresa se contraiga, muy probablemente hasta desaparecer. La empresa que canalice sus ganancias hacia el control de la contaminación no podrá recuperar sus pérdidas cobrando precios mayores a sus clientes. Los clientes no

263 (1980); Gerald P. Madden, "The Separation of Ownership from Control and Investment Performance", 34 *J. Econ. & Bus.,* 149 (1982); Sharon G. Levin y Stanford L. Levin, "Ownership and Control of Large Industrial Firms: Some New Evidence", 18 *Rev. Bus. & Econ. Research,* 37 (1983).

[56] Cotéjese el análisis de la inversión social a cargo de los inversionistas institucionales en el siguiente capítulo (véase después XVI.10).

[57] John W. Maxwell, Thomas P. Lyon y Steven C. Hackett, "Self-Regulation and Social Welfare: The Political Economy of Corporate Environmentalism", 43 *J. Law & Econ.,* 583 (2000).

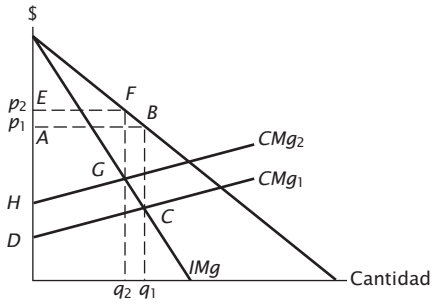
se benefician de tales gastos como clientes; más precisamente, se benefician por igual de tales gastos si compran el producto más barato de una empresa rival que no incurra en ellos. Por lo tanto, la empresa tendrá que sufragar los gastos del control de la contaminación enteramente con sus ganancias. Pero en un mercado competitivo no hay ganancias en un sentido económico, excepto como una consecuencia de la incertidumbre a corto plazo (donde los accionistas son los demandantes residuales de cualquier excedente de los ingresos corporativos sobre los costos).

Es cierto que, si una empresa tiene la curva de costo marginal habitual de pendiente positiva, podría ser capaz de continuar en la industria reduciendo su producción, aunque no minimizara sus costos. Pero no sobrevivirá indefinidamente, porque en su nivel de producción menor no podrá pagar a los dueños de los recursos escasos que utiliza (tierra, habilidades técnicas, etc.) una renta igual a la que esos dueños podrían obtener en otra parte; el monopsonio es raras veces un juego de largo plazo (véase antes X.13).

Las perspectivas para la responsabilidad social corporativa son sólo ligeramente más brillantes en los mercados monopólicos. Una empresa que no tiene rivales podrá trasladar a sus clientes una parte del costo del equipo de control de la contaminación, pero sólo una parte. Como se muestra en la gráfica XV.1, sus ganancias bajarán. Para el observador externo, el resultado es una reducción de las ganancias monopólicas (de *abcd* a *efgh*). Pero para el accionista es una pérdida. El precio de una acción es igual al valor presente de las ganancias futuras esperadas de esa acción.[58] Si la empresa tiene un monopolio con una perspectiva de ganancias monopólicas continuas de cierto nivel, el precio de la acción será mayor de lo que sería si se esperara un nivel de ganancias menor. Supongamos que la empresa decide incurrir en costos de control de la contaminación no esperados. Sus ganancias esperadas futuras son ahora menores, de modo que el precio de sus acciones bajará. Esto se sentirá como una pérdida para el accionista. En el caso habitual, el accionista no sabrá ni le importará que la corporación tenga ganancias monopólicas. Sólo le preocupa que el valor de sus propiedades haya declinado. Los administradores, a menos que el gasto en filantropía sea trivial en relación con los beneficios de la corporación, se resistirán a imponer tales consecuencias a sí mismos y a sus accionistas.[59]

[58] A menos que sea importante el valor de liquidación. Otra calificación de la que no necesitamos preocuparnos aquí, la volatilidad de las ganancias de la corporación, se discute más adelante, en XVI.1.

[59] Obsérvese la tensión existente entre diferentes metas sociales, aquí el abatimiento de la contaminación y la competencia. Entre más competitivo sea un mercado, más difícil será inducir a cualquier vendedor para que adopte un equipo costoso de control de la contaminación, porque si sus rivales no lo imitan —y el incentivo de éstos será la abstención— no podrá recuperar los costos del equipo. No sólo es un mercado monopólico más maleable al control de la contaminación, sino que la producción de contaminantes en tal mercado será normalmente menor. Cuando un mercado está monopolizado, la producción se reduce y, con ello, los subpro-



GRÁFICA XV.1

Y, si hay competencia para convertirse en un monopolista o para conservar el poder monopólico, un monopolista podría no tener ganancias con las cuales sufragar costos adicionales. El área *abcd* en la gráfica XV.1 podría no representar las ganancias monopólicas sino los costos fijos de haber obtenido el monopolio, en cuyo caso cualquier acción que disminuya esa área pondrá a la empresa en peligro de quiebra (véase antes IX.3).

Pero, por otra parte, ¿cómo explicar la práctica añeja de las corporaciones de hacer donaciones caritativas? Especialmente pero no sólo cuando se hacen en los lugares donde se ubican las plantas o las oficinas centrales de la corporación, son a menudo formas de relaciones públicas. Pueden ir encaminadas a conquistar el favor de los políticos y los grupos de interés, a hacerle publicidad a la corporación entre los clientes potenciales (por ejemplo, financiando eventos culturales), a rodearse de un aura de buena voluntad difusa, a ponerle un disfraz a la avaricia y a aplacar a los críticos. En ocasiones es posible lograr que las actividades comerciales se asemejen a la caridad. El costo marginal notoriamente bajo de los productos farmacéuticos hace que su venta en países pobres a precios muy bajos traiga una maximización de la ganancia, a condición de que se impida la reexportación de esos medicamentos a los países ricos. Los precios bajos pueden quedar representados como si estuvieran motivados por el altruismo.

La creciente globalización ha dado lugar a lo que podría denominarse "la caridad prodigada para hacer arrancar a la economía": donaciones dirigidas a construir la demanda futura de los productos del donador en los países pobres. Esto es particularmente cierto respecto a los productos que involucran externalidades de redes, de modo que el hecho de ser el primero en aparecer en un mercado puede darle a uno un empuje de entrada que será difícil superar. Entre más persona utilicen Windows, más valioso será Windows para otras personas. Y, una vez que se enganche con Windows, una persona pro-

ductos de la producción tales como la contaminación. Por supuesto, sin embargo, la elaboración de productos sustitutos aumentará y ello podría implicar una contaminación igual o mayor (así como la producción de equipo de control de la contaminación genera contaminación).

bablemente comprará programas de aplicaciones de Microsoft, como Office. En virtud de que el costo marginal de un sistema operativo es esencialmente cero, regalarlo puede significar la maximización de la ganancia al inducir a la compra de programas de aplicaciones que corren en él y crear un efecto de red; al mismo tiempo, el regalo puede mostrarse en los anuncios como un gesto filantrópico.

Empero, no hay que desconfiar de todas las obras de caridad de las corporaciones. Puede uno imaginarse situaciones en las que las actividades caritativas que una corporación realiza en forma desinteresada—donaciones que no incrementan la rentabilidad directa o indirectamente— podrían, no obstante, beneficiar a los accionistas y, por lo tanto, conllevar una maximización de la ganancia. La mayoría de los accionistas pueden abrigar el deseo de hacer algunas donaciones caritativas, pero creen que las corporaciones poseen más información que el individuo sobre las empresas de beneficencia que valen la pena. Las organizaciones de beneficencia difieren en la cantidad de dinero que gastan en su propia administración, como salarios y prestaciones para sus empleados, considerada en relación con la cantidad que destinan a los objetos reales de caridad. Las corporaciones se encuentran en una mejor posición que los individuos para determinar cuáles actividades caritativas son eficientes; si las cosas proceden así, entonces los accionistas pueden dar su consentimiento implícito a la decisión de sus corporaciones de realizar una cierta cantidad de actos caritativos. Pero no una gran cantidad. La caridad para una persona es la malicia diabólica para otra persona: el accionista que se opone al aborto por motivos religiosos se sentiría ofendido si su corporación contribuyera a la organización Planned Parenthood (Paternidad Planificada). Como resultado, las corporaciones evitan las actividades de beneficencia controvertidas, de modo que el asunto del consentimiento implícito acaba por reducirse a si el accionista vería con buenos ojos que la corporación hiciera una modesta contribución a actividades de beneficencia libres de controversia.

También existe una motivación fiscal detrás de dicho consentimiento. Si el accionista recibe un dividendo, la corporación habrá pagado el impuesto sobre la renta corporativo sobre el ingreso a partir del cual se paga el dividendo. Supongamos que la corporación y el accionista se hallan en el mismo grupo de 20%. La corporación gana 10 dólares, paga 2 de impuesto y le da al accionista 8. El accionista destina 8 dólares a actividades de beneficencia, lo que le cuesta 6.40, puesto que obtiene una deducción fiscal de 20%. Si quiere que la obra de caridad se quede con 10 dólares, tiene que hurgar en sus bolsillos en busca de otros 2, lo que le cuesta 1.60 (a causa de 20% de deducción), y, de este modo, el costo total para él de dar a la beneficencia 10 dólares es de 8. Si, en cambio, la corporación da los 10 dólares a la organización de beneficencia, un gasto deducible que por lo tanto le cuesta sólo 8 dólares, la organización de beneficencia recibe 10 dólares en vez de solamente, como antes, 8. El accionista pierde su deducción de 2 dólares, lo que significa que

el costo total de la transferencia para él es, como antes, de 8 dólares. Pero la corporación se encuentra en una mejor situación con el valor de 2 dólares, ya que así evita el impuesto fiscal corporativo sobre los 10 dólares en ingreso que dio a la organización de beneficencia. Y cualquier cosa que beneficie a la corporación beneficiará al accionista.

El último grito de la moda en materia de caridad corporativa circula bajo el nombre de "capitalismo creativo", y recibe el apoyo de Bill Gates, Warren Buffett y otros estadunidenses muy acaudalados.[60] Su rasgo distintivo es su afirmación de que la caridad corporativa es un buen negocio, aunque no por motivaciones egoístas. Gates, por ejemplo, sostiene que los jóvenes idealistas anhelarán trabajar para una compañía "capitalista de vena creativa", presumiblemente con un salario más bajo que si tuviesen que ser pagados para vencer sus escrúpulos sobre la perspectiva de trabajar para una maximizadora de la ganancia de cuño convencional. Dudamos de que haya muchas personas así. Pero hay algunas; y algunos de los proyectos caritativos en que Gates quisiera ver que se inmiscuyeran Microsoft y otras grandes compañías pueden resultar más interesantes que los proyectos de negocios puros acariciados por algunas empresas, y de esta manera pueden proporcionar un ingreso no pecuniario a los empleados merced al cual la empresa podría pagarles un ingreso más bajo. Existe aquí una analogía con el trabajo *pro bono* en algunos despachos jurídicos, el cual resulta atractivo para los jóvenes abogados porque suscita mayor interés que gran parte de su trabajo común y corriente, o por lo menos provee una variedad muy bienvenida. Aun así, si el "capitalismo creativo" reduce los costos de mano de obra lo suficiente para compensar el costo de la actividad caritativa, guarda consistencia con la maximización de la ganancia en sentido estricto.

Gates y Buffett creen que las corporaciones son más eficientes que las fundaciones y el gobierno, y, por ende, si la función de beneficencia se le transfiere al sector corporativo será llevada a cabo de manera más cumplida. (Buffett sugiere que 3% del ingreso derivado del impuesto sobre la renta corporativo podría asignarse a un fondo que "fuera administrado por algunos representantes del sector corporativo de los Estados Unidos a fin de que se empleara de modo inteligente para beneficio de la sociedad a largo plazo".) Pero las corporaciones son más eficientes sólo porque operan según la disciplina de la maximización de la ganancia, la cual se disipa si destinan recursos sustanciales a los proyectos caritativos. Existe incluso un peligro de que las corporaciones ejerzan presión sobre el gobierno para obligar a sus competidores de orientación no altruista a emprender por igual proyectos de beneficencia, a fin de impedir que esos rivales se apoderen del mercado.

[60] Véase *Creative Capitalism: A Conversation with Bill Gates, Warren Buffett, and Other Economic Leaders,* 63 (Michael Kinsley [comp.], 2008).

XV.15. La criminalidad corporativa

El tema más elemental de la responsabilidad social corporativa es el tocante a la obediencia que se debe a la ley. Estrictamente desde el punto de vista del accionista, podría parecer que los gerentes corporativos debieran obedecer a la ley solamente cuando los costos esperados de violarla sobrepasan a los beneficios esperados, por lo que los administradores tendrían el *deber* ante los accionistas de desobedecer a la ley, tal vez especialmente en los países donde la aplicación de la ley es muy débil (por ejemplo, un país que tuviera una ley contra el trabajo infantil pero fuera incapaz de hacerla cumplir). En este caso se trataría de una confrontación directa entre la ética y el bienestar del accionista. En vista de los beneficios sociales derivados del cumplimiento de la ley, el argumento que debería prevalecer de ordinario es el ético, de modo que a un accionista se le cerraría el paso si tratara de quejarse de que la administración corporativa, al no acertar a violar la ley aun cuando bien habría podido salirse con la suya, cometía una violación de su deber fiduciario para con sus accionistas. Otro argumento para el cumplimiento, que subyace a la prohibición de todo intento que hagan las empresas estadunidenses de involucrarse en sobornos incluso en países donde el soborno es de lo más común, es que la reducción del número de actos de esta naturaleza en esos países habrá de beneficiar a las compañías estadunidenses a largo plazo al inducir a los mercados de esos lares a ser más abiertos, lo que sería ventajoso para las empresas eficientes.

El hecho de que la violación cometida por los administradores a veces redunde en el interés de los accionistas nos lleva a exigir que se castigue a los primeros con la suficiente severidad para impedir que su penalización resulte contrarrestada por la recompensa que, es de esperarse, obtengan por haber generado las ganancias del accionista.

Una pregunta aparte es si debieran ser las corporaciones responsables por los delitos de sus administradores y otros empleados. Recuérdese que la doctrina de que el superior responde —que es responsable el empleador, independientemente de la ausencia de su falta personal, por los cuasidelitos cometidos por sus empleados dentro del campo de su empleo— se basa en el entendido de que la mayoría de los empleados no pueden pagar condenas monetarias sustanciales, de modo que la responsabilidad cuasidelictiva no tiene ningún efecto sobre sus incentivos. En virtud de que el derecho penal no se basa primordialmente en sanciones monetarias, de que la imposición de sanciones penales al empleador duplicaría las sanciones cuasidelictivas y de que las sanciones penales, por ser tan severas, pueden inducir a un cuidado excesivo, el derecho penal no ha adoptado la doctrina del *respondeat superior*. La excepción principal es la responsabilidad penal de las corporaciones. Si se comete o condona un delito, por lo menos ostensiblemente por cuenta

de la corporación, al nivel de los administradores o de la junta directiva, la corporación será penalmente responsable, lo que significa que la carga de la multa será soportada por los accionistas, quienes son análogos a los empleadores de las personas que realizaron efectivamente la acción.

La razón económica para la excepción es que, debido a que las corporaciones pueden ser sometidas sólo a sanciones financieras, a que son ya sea neutrales al riesgo o, si sienten alguna aversión por el riesgo, ésta es menor que en el caso de los individuos y a que el castigo corporativo lleva consigo escaso estigma (una corporación sólo puede actuar a través de individuos y hay una rotación constante de éstos), el castigo penal de la corporación es generalmente menos costoso para el castigado que el castigo de los individuos. Hay un peligro relativamente pequeño de que ello cause que los accionistas sean demasiado cuidadosos en la contratación, la supervisión y el despido de los directores (y, a través de la junta directiva, de los empleados administrativos). En estas circunstancias, podría ser eficiente la responsabilidad penal corporativa.

Supongamos por principio de cuentas que los administradores de la corporación son agentes perfectos de los accionistas, de modo que todo ingreso obtenido de la actividad delictiva llega a las manos de estos últimos. Entonces, si los accionistas no tienen ninguna responsabilidad financiera por un delito cometido por los administradores, tendrán grandes incentivos para contratar administradores que estén dispuestos a cometer delitos por cuenta de la corporación. Por supuesto, los accionistas tendrán que compensar a los administradores por los costos esperados del castigo penal. Pero, dados los limitantes existentes sobre la gravedad de las sanciones penales que se destacaron en el capítulo VII, los que truncan el costo esperado del castigo penal para los administradores, los accionistas pueden proceder de esta manera y tener todavía una ganancia esperada de la actividad delictiva corporativa. Supongamos ahora que los administradores no son agentes perfectos de la corporación: que, en efecto, utilizan sus posiciones corporativas para facilitar la actividad delictiva con vistas al enriquecimiento de ellos mismos. Sigue siendo cierto que la corporación ha proporcionado las instalaciones que los administradores están utilizando, y debieran otorgarse incentivos a sus propietarios para que seleccionaran y supervisaran a los administradores con mayor cuidado.

El verdadero enigma acerca de la responsabilidad penal corporativa, según parece, consiste en saber por qué debe haber una responsabilidad *penal*. La justificación económica del derecho penal es que el remedio cuasidelictivo óptimo es a menudo demasiado grande para ser reunido, ¿y cómo podría ser eso una consideración en el caso de una entidad que sólo puede ser sometida a sanciones monetarias? Sin embargo, las corporaciones no son infinitamente solventes, y dos de las técnicas fundamentales del derecho penal son plenamente aplicables incluso a una entidad que no puede ser castigada sino mediante una multa sin estigma: el uso de los recursos públicos para aumentar

la probabilidad del castigo por encima de lo que podría ser un nivel muy bajo debido a los esfuerzos realizados para ocultar la responsabilidad penal, y el castigo de la actividad preparatoria a fin de reducir la ganancia neta esperada del delito.

Pero, dado que el castigo penal corporativo es puramente monetario, no queda claro por qué debiera tener derecho la corporación a las elaboradas salvaguardias procedimentales del proceso penal. Como veremos en el capítulo XXII, tales salvaguardias tienen sentido económico sólo bajo el supuesto de que los castigos penales imponen grandes costos y no sólo una transferencia de dinero del delincuente al Estado.

La responsabilidad penal de la corporación ha sido criticada por varias razones.[61] Una de ellas es que tal responsabilidad no está integrada a la responsabilidad del cuasidelito, de modo que plantea un peligro sustancial de disuasión excesiva, dado que el acto delictivo de la corporación será a menudo también un cuasidelito que origina una responsabilidad sustancial de pago de daños. Puede otorgarse el pago de daños punitivos por el tipo de cuasidelito intencional que probablemente será también un delito, y las sanciones penales corporativas son puramente financieras, de modo que resulta difícil entender por qué debieran imponerse esas sanciones sin considerar la responsabilidad cuasidelictiva de la corporación. Otra crítica contra la responsabilidad penal corporativa consiste en que podría aumentar efectivamente la cantidad de delitos corporativos. Si los esfuerzos corporativos por impedir la comisión de delitos por parte de los empleados corporativos aumentan simultáneamente la probabilidad de que se descubran los delitos y, de ser descubiertos, de que se persigan luego, la corporación podría optar por reducir sus esfuerzos preventivos.

Este debate tiene implicaciones para la cuestión de si debieran ser las corporaciones responsables del pago de daños punitivos impuestos a sus empleados por su comportamiento que estuvo bajo la mira de su desempeño en su empleo.[62] Debido a que es probable que el empleado sea insolvente, es improbable que los daños punitivos se eviten a menos que la corporación sea responsable por ellos.

XV.16. La corporación cerrada

Hemos hecho hincapié en la corporación grande pública. Pero en los Estados Unidos gran parte de los negocios los realizan corporaciones cerradas. La corporación cerrada tiene pocos accionistas, la mayoría de los cuales, o incluso

[61] Véanse, por ejemplo, Jennifer Arlen, "The Potentially Perverse Effects of Corporate Criminal Liability", 23 *J. Legal Stud.,* 833 (1994); Daniel R. Fischel y Alan O. Sykes, "Corporate Crime", 25 *J. Legal Stud.,* 319 (1996).

[62] Véase A. Mitchell Polinsky y Steven Shavell, "Punitive Damages: An Economic Analysis", 111 *Harv. L. Rev.,* 869 (1998).

todos, están activos en la administración, y sus valores no se negocian públicamente, y a menudo no se pueden ni siquiera vender. El incentivo principal para formar una corporación cerrada, aparte de ciertas ventajas de reducción de los impuestos, es la obtención de una existencia perpetua y una responsabilidad limitada (aunque los acreedores grandes exigirán a menudo que los principales garanticen personalmente las deudas de la corporación).

Aunque muchos de los problemas que hemos considerado en relación con las corporaciones públicas no afligen a la corporación cerrada, ésta tiene sus propios problemas económicos, sobre todo el problema del monopolio bilateral. Supongamos que tres personas forman una corporación para la fabricación de software de computadora y cada una de ellas recibe un tercio de las acciones comunes de la corporación. Si alguno de ellos deseara más tarde dejar el negocio, probablemente no podría vender sus acciones a nadie que no sea uno de los otros accionistas (ya que cualquier extraño que comprara temería que los otros dos se unieran en su contra). Teniendo al próximo a partir parado en un barril de pólvora, los otros dos podrían ofrecerle muy poco por sus acciones. La solución consiste en celebrar un acuerdo entre los accionistas de adquisición o apoderamiento cuando se forme la corporación. Pero supongamos ahora que la corporación necesita más capital y, por lo tanto, desea vender acciones a una cuarta persona. Surgirá entonces esta interrogante: ¿cuál interés habrá de diluirse para hacerle lugar a este nuevo accionista? Cada uno de los accionistas existentes se abstendrá de ofrecer la dilución de su propio interés.

Podría pensarse que estos problemas se resuelven fácilmente por el principio de la regla de la mayoría. Pero, sobre todo cuando hay muy pocos electores, la regla de la mayoría puede ser muy inestable. Si dos de los accionistas de nuestra corporación cerrada hipotética tratan de coludirse contra el tercero, este último tendrá un incentivo para tratar de sonsacar a uno de los miembros de la coalición gobernante. Todo esto hará que los costos de negociación sean elevados y también que haya gran incertidumbre inicial. Una posible solución consiste en tener una regla de unanimidad para las transacciones que alteren la estructura corporativa, pero esto magnifica por supuesto el problema del monopolio bilateral.

Algo que la ley puede hacer y hace generalmente es permitir que los creadores de corporaciones cerradas se alejen libremente de la forma contractual estándar que es el derecho corporativo del Estado, a fin de que los socios redacten su propio convenio en una forma que no sería factible en una corporación muy grande, cuyos accionistas no pueden ser consultados eficazmente, en parte porque sus intereses en juego son demasiado pequeños para que les resulte rentable un estudio detenido de los detalles de la gobernanza corporativa. Otra posibilidad reconocida en la ley consiste en permitir que uno de los accionistas solicite la disolución de la corporación si hay un estancamiento que paralice las operaciones corporativas. Si se supone que los activos de la corpo-

ración valdrían menos después de la disolución, tal perspectiva aumentará el incentivo de los accionistas para negociar una salida del estancamiento. Por supuesto, como ocurre también en el divorcio —un análogo económico de la disolución de una corporación cerrada—, es muy importante que se condicione aceptar la petición a una compensación apropiada para los otros accionistas. De otro modo, el derecho de obtener la disolución se convertiría en un arma de negociación para cualquiera de los accionistas que tuviera menos que perder de la disolución.

Consideremos el caso In re Radom & Neidorff, Inc.[63] Radom y su cuñado tenían un exitoso negocio en el que los dos eran los únicos accionistas por partes iguales. El cuñado murió y su esposa, hermana de Radom, heredó sus acciones. Radom y su hermana no se entendieron. Aunque las reglas de la corporación requerían que ambos accionistas firmaran todos los cheques, ella se negó a firmar el cheque del sueldo de él, alegando que estaba ganando un sueldo excesivo. Aunque la corporación era rentable, el estancamiento entre sus accionistas le impedía declarar dividendos o aun pagar sus deudas.

Radom solicitó la disolución, pero el tribunal se negó a concederla. Si el tribunal hubiese concedido la petición de Radom, el resultado habría sido una expropiación efectiva del interés de su hermana, ya que Radom, el socio activo, podría haber continuado fácilmente el negocio bajo un nombre nuevo. Eso sería como otorgar un divorcio unilateral a un hombre cuya esposa acabara de sostenerlo durante toda su educación en la escuela de medicina (véase antes V.3). El tribunal podría haber condicionado la disolución a que Radom le pagara a su hermana una parte igual del valor de la corporación como un negocio en marcha con buen funcionamiento (como un acuerdo de propiedad en un divorcio), en lugar de dividir simplemente con ella sus activos tangibles. Pero es posible que la negativa de la petición sirviera igualmente, ya que presionaba a Radom (aunque no sabemos con qué efectividad) para que le ofreciera a su hermana un precio mayor por sus acciones.

Este análisis sugiere una razón para la regla de que una asociación se disuelve a la muerte de cualquiera de los socios: Si la asociación continuara, los socios sobrevivientes tendrían a la viuda del finado parada sobre un barril, como en el caso Radom. La disolución le otorga a ella cierto poder de negociación.

XV.17. Una revisión de la regulación del servicio público

Al emplear términos como “deuda” y “capital social” en este capítulo, nos hemos apegado estrechamente a sus significados legales, como se hace de ordinario en este libro, y ello puede oscurecer interesantes cuestiones económicas.

[63] 119 N.E. 2d, 563 (N.Y., 1954).

Veamos la empresa de servicio público que discutimos en el capítulo XIII. Formalmente, es una corporación convencional con accionistas y tenedores de bonos. Pero, en la medida en que la regulación alcanza su objetivo formal de fijar el precio del servicio sobre la base de un margen de ganancia sobre el costo, altera las relaciones económicas reales existentes dentro de la empresa. El accionista ya no es el reclamante residual. Los cambios de fortuna de la compañía se registran en el precio de su producto o servicio, antes que en el precio de sus acciones. Por lo tanto, el consumidor es el verdadero tenedor del capital. Esto sugiere la objeción económica fundamental contra el margen de ganancia sobre el costo. El consumidor se encuentra en una posición peor aún que la del accionista para supervisar a los administradores, ya que no tiene poder de votación y no puede vender sus acciones. Pero, en un sistema de margen de ganancia sobre el costo, el consumidor es la única persona que tiene un incentivo económico para supervisar a los administradores, más allá del incentivo limitado que tiene todo acreedor a largo plazo para vigilar al deudor.

Lecturas recomendadas

Ayotte, Kenneth M., y Edward R. Morrison, "Creditor Control and Conflict in Chapter 11", 1 *J. Legal Analysis,* 511 (2009).

Ayres, Ian, y Joe Bankman, "Substitutes for Insider Trading", *54 Stan. L. Rev.,* 235 (2001).

Baird, Douglas G., "Bankruptcy's Uncontested Axioms", 108 *Yale L. J.,* 573 (1998).

———, "Fraudulent Conveyances, Agency Costs, and Leveraged Buyouts", 20 *J. Legal Stud.,* 1 (1991).

———, y Anthony J. Casey, "No Exit? Withdrawal Rights and the Law of Corporate Reorganizations", 113 *Colum. L. Rev.,* 1 (2013).

Bebchuk, Lucian Arye, y Jesse M. Fried, *Pay without Performance: The Unfulfilled Promise of Executive Compensation* (2004).

———, "The Uneasy Case for the Priority of Secured Claims in Bankruptcy", 105 *Yale L. J.,* 857 (1996).

Corporate Bankruptcy: Economic and Legal Perspectives (Jagdeep S. Bhandari y Lawrence A. Weiss [comps.], 1996).

Corporate Law and Economic Analysis (Lucian Arye Bebchuk [comp.], 1990).

Crémer, Jacques, "Corporate Culture and Shared Knowledge", 2 *Industrial & Corporate Change,* 351 (1993).

Easterbrook, Frank H., y Daniel R. Fischel, *The Economic Structure of Corporate Law* (1991).

Economics of Bankruptcy (2 vols., Edward R. Morrison [comp.], 2012).

Engle, David L., "An Approach to Corporate Social Responsibility", 32 *Stan. L. Rev.*, 1 (1979).

Gordon, Jeffrey N., "Shareholder Initiative: A Social Choice and Game Theoretic Approach to Corporate Law", 60 *U. Cinc. L. Rev.,* 347 (1991).
Hansmann, Henry, *The Ownership of Enterprise* (1996).
Jensen, Michael C., y William H. Meckling, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure", 3 *J. Fin. Econ*., 305 (1976).
Organization Theory: From Chester Barnard to the Present and Beyond (Oliver E. Williamson [comp.], 1990).
Posner, Richard A., "From the New Institutional Economics to Organization Economics: With Applications to Corporate Governance, Government Agencies, and Legal Institutions", 6 *J. Institutional Econ.,* 1 (2010).
Prendergast, Candice, "The Provision of Incentives in Firms", 37 *J. Econ. Lit*., 7 (1999).
Roberts, John, *The Modern Firm: Organizational Design for Performance and Growth* (2004).
Shleifer, Andrei, y Robert W. Vishny, "A Survey of Corporate Governance", 52 *J. Fin.,* 737 (1997).
Ulen, Thomas S., "The Coasean Firm in Law and Economics", 18 *J. Corp. L.*, 301 (1993).
White, Michelle J., "Corporate Bankruptcy as a Filtering Device: Chapter 11 Reorganizations and Out-of-Court Debt Restructurings", 10 *J. Law, Econ. & Organization,* 268 (1994).
———, "Bankruptcy: Past Puzzles, Recent Reform, and the Mortgage Crisis", 11 *Am. Law & Econ. Rev.,* 1 (2009).

Problemas

1. Supongamos que cuando el correo era todavía un departamento ejecutivo se le pidiera diseñar una forma organizacional apropiada para el Servicio Postal. ¿Propondría una corporación privada? ¿Una corporación pública? ¿Sería cualquiera de esas formas un adelanto sobre el departamento ejecutivo?

2. La Corporación de Satélites de Comunicaciones (Comsat) es una corporación privada, pero el presidente de los Estados Unidos nombra a varios de los directores (una minoría). ¿Sería de esperarse que la empresa se comportara de manera diferente a la de otras corporaciones privadas por la presencia de estos directores públicos?

3. Según la ley alemana, los trabajadores tienen derecho a elegir a algunos de los miembros del equivalente alemán (en términos aproximados) de nuestra junta directiva. ¿Diría que ésta es una innovación que convendría hacer en nuestra forma corporativa? Consúltese a Detlev F. Vagts, "Reforming the 'Modern' Corporation: Perspectives from the German", 80 *Harv. L. Rev*., 23 (1966). ¿Dependerá su respuesta de lo que maximicen los líderes sindicales?

(véase el problema 6). Supongamos que se entregaran a los trabajadores acciones del capital corporativo: ¿habría alguna diferencia si fueran acciones de la compañía para la que trabajaban o acciones de alguna otra compañía?

4. Discuta esta proposición: El derecho de las corporaciones debiera ser enteramente permisivo en lo tocante al reparto del poder, en la escritura constitutiva, entre accionistas y directores.

5. ¿Cómo podría beneficiarse un administrador de adoptar una meta corporativa de maximización de las ventas? ¿Depende su respuesta de que el mercado de administradores sea competitivo? ¿De que los compradores de ese mercado estuviesen bien informados?

6. ¿Esperaría que el problema de que los administradores sustituyen las metas de la organización por sus propias metas personales fuese más agudo en una universidad, un sindicato o una dependencia gubernamental reguladora que en una empresa comercial? ¿Puede argüirse que las metas personales de los administradores de cualquier clase de organización coinciden generalmente con la meta de maximización del valor presente de la organización?

7. Para aplicar la regla contra la negociación de las personas que están dentro de la organización, ¿debieran incluirse sólo a los funcionarios corporativos, sus familias y las personas a quienes venden información confidencial, o debiera incluirse también a quienquiera que posea en efecto información confidencial? ¿Por qué podría ser en extremo costosa la definición extensa, muy aparte de los costos administrativos y de ejecución?

8. La Ley Price-Anderson, 142 U.S.C., 2210, impone en efecto una limitación (de varios cientos de millones de dólares) a la responsabilidad cuasidelictiva de cualquier fabricante o propietario de un reactor nuclear por las consecuencias de un accidente del reactor. ¿Cuál es la justificación económica de esta limitación?

9. ¿Esperaría que la discriminación racial fuese más común en una corporación pública o en una cerrada? ¿En una empresa regulada o en una no regulada? Supongamos que la política pública es neutral acerca de la discriminación privada. Si no lo tiene claro, consúltese después XXVIII.9.

10. Supongamos que hay una primera y una segunda hipotecas sobre un inmueble. Si el deudor hipotecario no paga, ¿quién debiera tener el derecho de demandar desde el punto de vista de la eficiencia económica: el primer acreedor hipotecario o el segundo?

11. Un ferrocarril salió de la quiebra en 1896 con una estructura de capital alterada. Los acreedores de la primera hipoteca recibieron el total de las acciones comunes. Los acreedores de la segunda hipoteca y los accionistas recibieron dos clases de títulos (un "título" es un bono que no está asegurado por ningún gravamen sobre los inmuebles del emisor): clase A y clase B. La clase A, emitida a cambio de una aportación de fondos nuevos a la corporación reorganizada, consiste en 600 títulos con un valor nominal de 1 000 dólares cada uno. La clase B consiste en 7 000 títulos con el mismo valor nominal.

Los títulos no tienen fecha de vencimiento ni un derecho de interés fijo. Los certificados de títulos proveen que, si la junta directiva del ferrocarril declara dividendos, los accionistas (es decir, los accionistas nuevos, los acreedores de la primera hipoteca antes de la reorganización) y los tenedores de los títulos de la clase A tendrán derecho a la porción de los dividendos que se requiera para dar a estas dos clases de tenedores de valores 5% del valor nominal de sus valores (valor nominal, en el caso de las acciones comunes). Los tenedores de los títulos de la clase B tienen derecho al resto. Los certificados proveen además que, en caso de una venta o una reorganización del ferrocarril, los primeros 600 000 dólares del monto de la venta se entregarán a los tenedores de los títulos de la clase A, los siguientes 2.5 millones de dólares a los accionistas, y el resto —o los primeros 7 millones de dólares del resto— a los tenedores de los títulos de la clase B. Sin embargo, los certificados son ambiguos. La primera oración dice que el monto nominal de los títulos (7 millones de dólares para los 7 000 títulos de 1 000 dólares) será pagadero sólo en el caso de una venta o una reorganización, pero la última oración dice que cualesquiera fondos netos restantes después del pago de los 3.1 millones de dólares a los otros tenedores de valores se repartirán a prorrata entre los tenedores de los títulos de la clase B.

Casi un siglo después de la reorganización, el ferrocarril se vendió por bastante más que 10.1 millones de dólares (la suma de los derechos establecidos en los certificados). ¿Quién debiera recibir el excedente: los propietarios de las acciones comunes, como en el caso habitual de la venta de un negocio, o los tenedores de títulos de la clase B, en los términos de la última oración del certificado? ¿Cuál resultado tiene más sentido desde un punto de vista económico? Véase Eliasen *vs.* Itel Corp., 82 F.3d, 731 (7th Cir., 1996).

12. El capítulo 9 del Código de Quiebras permite que los municipios se declaren en quiebra y presenten un plan de reorganización análogo al permitido por el capítulo 11. (Sólo se permite la quiebra voluntaria; los acreedores del municipio no pueden forzarlo a la quiebra.) Una cuestión controversial consiste en saber si el tribunal de quiebras está facultado, o debiera estarlo, para condicionar la confirmación del plan a que el municipio aumente los impuestos para ayudarse a pagar a sus acreedores. Evalúe esta cuestión desde el punto de vista de la economía, considerando: *1)* si las quiebras municipales serían más o menos probables en el caso de que el tribunal de quiebras tuviera esta facultad; *2)* si es probable que los residentes o los acreedores soporten el riesgo del impago municipal de manera más eficiente; *3)* cuál grupo tiene probabilidades de ser un mejor monitor de las deudas de un municipio. Véanse opiniones contrarias en Michael W. McConnell y Randal C. Picker, "When Cities Go Broke: A Conceptual Introduction to Municipal Bankruptcy", 60 *U. Chi. L. Rev.,* 425 (1993), por una parte, y Kevin A. Kordana, "Tax Increases in Municipal Bankruptcies", 83 *Va. L. Rev.,* 1035 (1997), por la otra.

13. ¿Debiera permitirse que un deudor y sus acreedores convinieran, en el momento del otorgamiento del crédito, en que el deudor no estará sujeto al Código de Quiebras? Véase Steven L. Schwarcz, "Rethinking Freedom of Contract: A Bankruptcy Paradigm", 77 *Texas L. Rev*., 515 (1999).

14. Se permite que un deudor, en una quiebra de consumidor, "reafirme" una deuda, es decir, convenga en que la deuda (de ordinario una deuda asegurada con un gravamen sobre la propiedad personal, tal como una joya) no será pagada en la quiebra. ¿Por qué haría eso un deudor? Véase Barry Adler, Ben Polak y Alan Schwartz, "Regulating Consumer Bankruptcy: A Theoretical Inquiry", 29 *J. Legal Stud.,* 585, 590 (2000).

15. ¿En qué grado piensa que las juntas directivas de organizaciones sin fines de lucro, como universidades, hospitales de beneficencia y museos, difieren y, desde un ángulo económico, debieran diferir de las juntas directivas de corporaciones de negocios? Véase Katherine O'Regan y Sharon M. Oster, "Does the Structure and Composition of the Board Matter? The Case of Non-Profit Organizations", 21 *J. Law, Econ. & Organization,* 205 (2005).

XVI. LOS MERCADOS FINANCIEROS

XVI.1. El diseño de la cartera

En este capítulo, nuestro enfoque se traslada de la corporación misma al accionista, a los administradores de carteras de inversión —administradores fiduciarios, planes de pensiones, bancos, fondos mutuos de inversión, compañías de seguros e inversionistas individuales— y al sistema financiero en general de bancos y reguladores financieros. El colapso financiero de 2008 —que desencadenó la mayor recesión financiera, en los Estados Unidos y en el mundo entero, desde la Gran Depresión del decenio de 1930, y que nos siguió acompañando cinco años después— ha sembrado escepticismo acerca de la teoría de las finanzas, y esta reacción merece una consideración seria.

Un lugar donde se puede comenzar el análisis es el concerniente al mercado de valores corporativos, en particular el mercado de las acciones comunes de la empresa. Un valor tiene dos dimensiones: el riesgo y el rendimiento esperado. El rendimiento esperado se obtiene multiplicando cada rendimiento posible por la probabilidad de que sea el rendimiento efectivo, y se suman luego los resultados de la multiplicación. Por ejemplo, si hay una probabilidad de 50% de que una acción particular que se vende hoy por 10 dólares valga 12 dentro de un año, una probabilidad de 40% de que valga 15 y una probabilidad de 10% de que no valga nada, su rendimiento esperado es 2 dólares [(0.5 × 2 dólares) + (0.4 × 5 dólares) – (0.1 × 10 dólares)].[1]

Aunque el rendimiento esperado de una probabilidad de 100% de obtener 10 dólares es el mismo (10 dólares) que el rendimiento esperado de una probabilidad de 50% de obtener 20 dólares o de una probabilidad de 1% de obtener 1 000 dólares, sabemos que la gente no es indiferente a las diversas formas en que se combinan la incertidumbre y los resultados para obtener el mismo rendimiento esperado. Al escoger entre los valores que tienen rendimientos esperados idénticos, el inversionista que siente aversión por el riesgo escogerá el que tenga la incertidumbre menor, a menos que bajen los precios de los otros valores, aumentando así sus rendimientos esperados, hasta el punto de que se siente adecuadamente compensado por asumir un riesgo mayor.

El predominio de la aversión al riesgo en la inversión se ejemplifica por la tasa de rendimiento normalmente menor de los bonos por comparación con las acciones comunes de la misma compañía. Supongamos que el rendi-

[1] El valor esperado es $12 [(0.5 × $12) + (0.4 × $15) + (0.1 × $0)], y el precio corriente es de 10 dólares. A fin de simplificar el análisis, se supone que no se pagarán dividendos. El rendimiento esperado de una acción incluye, por supuesto, tanto la apreciación como los dividendos.

miento esperado (dividendos más apreciación) de las acciones comunes de una compañía es de 10%. Los inversionistas que son neutrales al riesgo demandarían 10% de interés también en los bonos de la compañía. Aunque es menos riesgoso ser tenedor de bonos, porque tiene el colchón de los accionistas, quienes tendrían que ser descartados por completo antes de que él pudiera perder sus intereses, su seguridad adicional se ve contrarrestada en el sentido del rendimiento esperado por el hecho de que no puede ganar más que la tasa de interés especificada en el bono. La diferencia entre la tasa de interés del bono de una compañía y el rendimiento esperado (mayor) de los dueños de las acciones comunes es una compensación a los accionistas por el riesgo mayor que asumen.[2]

Se sigue de aquí que debiera haber también una diferencia sistemática entre los rendimientos esperados de acciones comunes que difieren en cuanto a su riesgo; pero esta observación está sujeta a una calificación importante. Supongamos que los rendimientos esperados por acción de dos acciones (A y B) son iguales, 2 dólares, pero el rendimiento esperado de A combina una probabilidad de 50% de cero rendimiento con una probabilidad de 50% de un rendimiento de 4 dólares, mientras que el rendimiento esperado de B combina una probabilidad de 50% de un rendimiento de –6 dólares con una probabilidad igual de un rendimiento de 10 dólares. B es la acción más riesgosa. Supongamos que hay una tercera acción (C) que, como B, tiene un rendimiento esperado de 2 dólares resultante de la combinación de una probabilidad de 50% de un rendimiento de –6 dólares con la misma probabilidad de un rendimiento de 10, sólo que las fortunas de C y de B son recíprocas, de modo que cuando a B le va bien le va mal a C y viceversa.[3] Entonces, una *cartera* compuesta de B y C será menos riesgosa que una compuesta sólo de A, a pesar de que A, considerada aisladamente, sea menos riesgosa que B o C. El inversionista no insistirá en un premio de riesgo por mantener a B y C en su cartera. Sus riesgos se cancelan; la cartera misma está libre de riesgo.

Esto ilustra la observación fundamental de que el diseño de la cartera puede alterar las características de riesgo de los valores considerados de manera individual. Y, en un mundo donde los riesgos de acciones comunes diferentes estuvieran negativamente correlacionados, como en el ejemplo anterior, habría pocos premios por el riesgo diferencial, si acaso, entre las acciones comunes. Menos obviamente, esto se aplicaría también si los riesgos de las

[2] Pero existe un riesgo que los tenedores de bonos asumen mientras que los accionistas no lo hacen (en la misma medida): el riesgo de un cambio no previsto en la tasa de inflación. El rendimiento de los tenedores de bonos es un porcentaje fijo de la cantidad aparente del bono. Dado que la cantidad aparente no se transforma con la inflación, tampoco lo hace el rendimiento de los tenedores de bonos.

[3] Es decir, hay una probabilidad de 50% de que B genere un rendimiento de –6 dólares y C un rendimiento de 10 dólares, y una probabilidad de 50% de que B genere un rendimiento de 10 dólares y C un rendimiento de –6 dólares.

acciones comunes, en lugar de estar correlacionados negativamente, fueran no correlacionados, esto es, aleatorios; porque, en una cartera integrada por muchas acciones comunes diferentes, los riesgos de las acciones de la cartera distribuidos aleatoriamente tenderían a cancelarse. De un modo análogo, obsérvese que, si bien el riesgo de cada vida en este país no es insignificante, la tasa de mortalidad del país —el funcionamiento de la "cartera" de todos los individuos— es extremadamente estable. En efecto, es mucho más estable que el mercado de valores. Esto sugiere que los riesgos de diferentes acciones comunes no se correlacionan negativamente ni en forma aleatoria, sino que en efecto tienen una fuerte correlación positiva, lo que a su vez vuelve necesario que en el diseño de la cartera se distinga entre dos componentes del riesgo. Uno es el componente que se correlaciona positivamente con el riesgo del conjunto de las acciones, el mercado, y que por lo tanto no puede eliminarse tan sólo añadiendo más y más valores. El otro componente es el riesgo que se correlaciona negativamente, o no se correlaciona, con el riesgo del mercado en conjunto, de modo que puede ser eliminado mediante la diversificación. La diversificación es una meta importante del diseño de la cartera porque nos permite librarnos de una forma del riesgo que no está compensada (precisamente porque puede eliminarse a escaso costo mediante la diversificación) y es, por ende, una pérdida de peso muerto para el inversionista que siente aversión por el riesgo. Pero la diversificación no elimina todo el riesgo; como hemos visto, una parte del riesgo no se puede diversificar, y a fin de soportar ese riesgo insistirá el inversionista (que suponemos siente aversión por el riesgo) en una compensación. (Además, al difundir el riesgo, la diversificación impone el riesgo, aunque usualmente de forma tolerable: si usted pertenece a un fondo de seguros, les impone riesgos a los otros miembros del fondo.) En virtud de que el riesgo sistemático —el componente del riesgo que se correlaciona positivamente con el riesgo del mercado en conjunto— es también un riesgo compensado, el administrador de la cartera que desee reducirlo deberá estar dispuesto a pagar un precio bajo la forma de un rendimiento esperado menor.[4]

Consideremos un valor que aumenta en promedio 10% cuando el mercado aumenta 10% y baja 10% cuando el mercado baja 10%. Su riesgo sistemático sería igual al del mercado. En el lenguaje de las finanzas, tendría un beta de 1, siendo beta el riesgo del valor en relación con el del mercado en conjunto. Si el valor aumentara 20% cuando el mercado aumentara 10% y bajara

[4] En el terreno empírico se ha descubierto que las acciones que difieren en cuanto al riesgo sistemático difieren también en cuanto al rendimiento esperado, y se ha hallado asimismo que la correlación existente entre el riesgo sistemático y el rendimiento es positiva, como sería de esperarse. Las pruebas se resumen en James H. Lorie y Mary T. Hamilton, *The Stock Market: Theories and Evidence*, caps. 11-12 (1973), pero siguen siendo controversiales. Compárese Eugene F. Fama y Kenneth R. French, "Common Risk Factors in the Returns on Stocks and Bonds", 33 *J. Financial Econ.,* 3 (1993), con Fischer Black, "Beta and Return", 20 *J. Portfolio Mgmt.,* 8 (otoño de 1993).

20% cuando el mercado bajara 10%, su beta sería 2, y si aumentara sólo 5% cuando el mercado aumentara 10% y bajara 5% cuando el mercado bajara 10%, su beta sería 0.5.[5] Por supuesto, ningún valor se mueve perfectamente al paso del mercado. Pero no nos interesa la porción de la varianza del valor que no se correlaciona con el movimiento del mercado en conjunto, que es diversificable. Una acción con un beta elevado debiera tener un rendimiento esperado elevado, mientras que una acción que tenga un riesgo *global* mayor pero un beta bajo —una acción que se mueva mucho pero no al paso del mercado en conjunto— debiera tener un rendimiento esperado bajo.

¿Por qué difieren las acciones en su grado de riesgo, ya sea sistemático o aleatorio (diversificable)? Dos factores son particularmente importantes. Uno, que examinaremos en la sección siguiente de este capítulo, es el grado en el que las actividades de la corporación están diversificadas. El otro es la razón de deuda-capital accionario (apalancamiento) de ésta. Si una compañía tiene una razón elevada de deuda de largo plazo a capital accionario en su estructura de capital, cualquier declinación de los ingresos brutos de la empresa golpeará a los accionistas más duro porque la deuda es un gasto fijo; no desciende cuando los ingresos descienden, a diferencia de los costos variables, que bajan cuando la producción está a la baja (y con ella los rendimientos). Las acciones de tal empresa tenderán a tener un beta elevado porque cualquier evento que afecte adversamente al mundo de los negocios (y, por ende, al mercado de valores) en conjunto, tal como una baja no esperada de la demanda agregada, afectará a esta compañía más adversamente aún. En cambio, si sube el mercado en conjunto (digamos a causa de un aumento de la demanda agregada), esta compañía tenderá a encontrarse mejor que las empresas que tienen cantidades de deuda promedio. Como un ejemplo del riesgo diversificable, consideremos ahora una compañía que se dedica a la búsqueda de uranio. Su fortuna tenderá a aumentar y disminuir con su éxito o fracaso en la localización de depósitos de uranio, que es un proceso incierto pero que con toda probabilidad no tenderá a relacionarse sistemáticamente con los movimientos del mercado de valores en conjunto. Puede constituir una reserva de acciones muy riesgosa, pero su beta no podría ser mayor al promedio.

El análisis anterior sugiere que, si dejamos de lado los costos administrativos de la diversificación, la cartera de un inversionista debiera diversificarse tanto como se pueda a fin de eliminar el riesgo no compensado. Una medida de la diversificación de una cartera es su correlación con algún índice de base amplia de las oportunidades de inversión, tal como el Standard & Poor's 500. Pero, en virtud de que los movimientos de una cartera integrada por sólo 32 acciones (cuidadosamente seleccionadas) tendrían una correlación de 95% con los movimientos del S&P 500, se supone a veces que no tiene caso

[5] James H. Lorie, "Diversification: Old and New", 1 *J. Portfolio Mgmt*., 25, 28 (invierno de 1975).

mantener una cartera mayor, ya no digamos una que incluiría 250 o 500 acciones. Esto es incorrecto. Por ejemplo, aunque una cartera integrada por 50 acciones tendría un coeficiente de correlación muy por arriba de 90%, su rendimiento esperado sería un intervalo de 4.5 puntos porcentuales a cada lado del rendimiento esperado del S&P 500, de modo que si el S&P 500 aumentara 10% a lo largo del año, se esperaría que la cartera de 50 acciones aumentara entre 5.5 y 14.5%. Aun una cartera integrada por 100 acciones diferiría a menudo hasta 3% del desempeño del S&P 500; se necesita una cartera de 200 acciones para reducir esta cifra a 1%.[6] Y ni siquiera el S&P 500 está completamente diversificado; es una muestra de sólo un tipo de activo que se negocia en una sola bolsa de valores.

Otra implicación del análisis es que el administrador de la cartera, por su elección del beta (riesgo no diversificable) de la cartera, determinará el rendimiento esperado del inversionista. Por lo tanto, el administrador de la cartera puede variar tal rendimiento sin tratar de atinarles a los ganadores. El método más eficiente para alcanzar la combinación deseada de riesgo y rendimiento consiste en ajustar las proporciones en las que se incluyan activos relativamente libres de riesgo en la cartera o se utilice dinero prestado para aumentar las propiedades de la cartera. Consideremos como un punto de referencia la cartera de mercado hipotética integrada por todas las acciones negociadas en el mercado de valores,[7] ponderadas por el valor de mercado de cada compañía cuyas acciones están representadas. Por supuesto, el beta de la cartera de mercado es 1, y es probable que su rendimiento esperado ahora sea de alrededor de 9%.[8] Supongamos que el administrador de la cartera está dispuesto a aceptar un rendimiento menor a cambio de una volatilidad menor, digamos la mitad de la del mercado en conjunto. Es decir, desea una cartera que tendrá un beta de 0.5. ¿Cómo podrá lograr esto? Podría simplemente rechazar las acciones más riesgosas de la cartera del mercado hasta que el beta promedio de las acciones restantes fuese de sólo 0.5. Pero su cartera se volvería menos diversificada en el proceso y, como hemos señalado varias veces, los riesgos evitables por una diversificación mayor no están compensados. La alternativa consiste en añadir a la cartera suficientes bonos corporativos o gubernamentales, u otros valores de ingreso fijo, con sus betas típicamente bajos, para empujar al beta promedio de la cartera a la baja hacia el 0.5. Esto

[6] *Idem.* Estas cifras dependen de la manera precisa como se construya la muestra. Quizá pudieran obtenerse mejores resultados utilizando muestras más pequeñas.

[7] Por supuesto, hay más de un mercado de valores. Pero dado que, en términos de su valor, cerca de dos tercios del total de las acciones negociadas públicamente en los Estados Unidos se encuentran listadas en la Bolsa de Valores de Nueva York (BVNY), esa bolsa provee una buena aproximación al mercado de valores (de los Estados Unidos) en conjunto; por lo tanto, consideraremos que nuestra cartera de mercado hipotética se limita a las acciones de la BVNY.

[8] Históricamente, el rendimiento esperado para las acciones comunes ha tenido un promedio de alrededor de 3 puntos porcentuales por encima de la tasa de bonos corporativos a largo plazo. Eugene F. Fama & Kenneth R. French, "The Equity Premium", 57 *J. Fin.,* 637 (2002).

tiene la ventaja de no reducir la diversificación del componente de acciones comunes de la cartera. Es una mejor estrategia que remplazar las acciones comunes, en la cartera, con bonos que tienen un beta promedio de 0.5, porque la cartera resultante quedaría demasiado poco diversificada; por ejemplo, estaría mucho más expuesta al riesgo de un cambio inesperado en la tasa inflacionaria que una cartera que incluyera acciones comunes, cuyas ganancias no están fijas en términos de dólares nominales. Una observación relacionada es que los bonos añadidos a una cartera de acciones comunes, a fin de reducir el beta de la cartera, debieran seleccionarse con vistas a mantener la diversificación global de la cartera. No se lograría tal objetivo con una cartera consistente en un millón de dólares en las acciones de un fondo de mercado (es decir, un fondo que mantiene la cartera del mercado o alguna aproximación razonable) y un millón de dólares en los bonos de una compañía.

El inversionista que desee un rendimiento esperado mayor que el del mercado en conjunto podrá obtenerlo pidiendo dinero prestado a fin de comprar valores adicionales para la cartera. Supongamos que un inversionista de 500 000 dólares de su propio dinero toma prestados otros 500 000, lo que le da un millón de dólares en activos que invierte en un fondo que tiene la cartera del mercado; este inversionista paga un interés de 6% sobre el préstamo; y el rendimiento esperado de la cartera del mercado es de 9%. Su tasa de rendimiento esperado, que es igual al rendimiento esperado de la cartera (0.09 × un millón = 90 000 dólares) menos sus costos de intereses (0.06 × 500 000 = 30 000 dólares), dividido por su inversión personal, será de 12% (90 000 – 30 000 = 60 000 dólares dividido por 500 000 dólares = 0.12). Esto es mayor que la tasa del mercado; pero el beta de la cartera es mayor que el beta del mercado. Si, por ejemplo, el mercado declinara 10%, la cartera valdría sólo 900 000 dólares (podemos omitir el costo de intereses del dinero que se toma prestado), y, dado que el inversionista debería todavía 500 000 dólares al prestamista, sus activos netos serían sólo 400 000 dólares, 20% menos que antes de la declinación del mercado. El beta de la cartera es entonces 2.

XVI.2. Diversificación, apalancamiento y la razón de deuda-capital accionario

El análisis de la sección anterior nos ayuda a evaluar el argumento de que la existencia de corporaciones conglomeradas —corporaciones cuyas divisiones o subsidiarias operan en mercados no relacionados, condenando así aparentemente a la corporación a las deseconomías de la especialización insuficiente— se debe al hecho de que la empresa conglomerada supera ciertas ineficiencias de las estructuras financieras de muchas empresas. Los conglomerados están a menudo muy diversificados, y se ha sostenido que una empresa muy diversificada, al reducir el riesgo para los inversionistas, permite que se adquiera capital a un costo menor que si la empresa no estuviera di-

versificada. Este argumento no toma en cuenta el hecho de que el accionista cuya cartera está integrada por acciones de empresas que no están diversificadas pero operan en mercados no relacionados disfruta los mismos beneficios de diversificación que tendría si esas empresas estuvieran consolidadas en un solo conglomerado. Lo importante para un inversionista es la diversificación de su cartera, no que las acciones estén diversificadas.

Una segunda argumentación es que muchas empresas compradas por las corporaciones conglomeradas tienen razones bajas de deuda a capital accionario, y que el conglomerado, al aumentar esa razón, puede aumentar el rendimiento para los accionistas. Esta argumentación se suma con frecuencia a la primera en un esfuerzo de persuadirnos de que la diversificación del conglomerado le permite minimizar los riesgos asociados a las razones de deuda-capital accionario elevadas. Pero es improbable que el valor de las acciones de los accionistas pueda aumentar modificando la razón de deuda a capital accionario.[9]

Consideremos una empresa hipotética que tiene 1 millón de acciones comunes vigentes, sin deuda, y un ingreso de operación neto sostenido de 3 millones de dólares anuales. Dado que el valor de una empresa es el valor presente descontado de su ingreso esperado, el valor de nuestra empresa hipotética será un múltiplo de 3 millones de dólares (el múltiplo está determinado por cosas tales como el beta de las reservas y la demanda de activos de inversión en general —y también por las ganancias futuras estimadas de la firma—, pero ese determinante se elimina por el supuesto en el ejemplo); y supongamos que es 10. Entonces, la empresa valdrá 30 millones de dólares y el precio de una acción de reserva de la empresa será de 30 dólares, las ganancias por acción de 3 dólares y la razón de precio a ganancias será de 10. Supongamos ahora que la empresa toma prestados 15 millones de dólares con una tasa de interés de 6% y utiliza este dinero para recomprar la mitad de sus acciones comunes vigentes, que luego retira. El ingreso de operación neto no ha cambiado, pero ahora hay un gasto de intereses de 900 000 dólares. Esto reduce las ganancias de los accionistas a 2.1 millones de dólares, o sea, 4.20 dólares por acción porque el número de acciones ha bajado de 1 millón a 500 000. Si no cambia la razón de precio a ganancias, cada acción valdrá 42 dólares y el capital accionario agregado de los accionistas valdrá 21 millones de dólares. Por lo tanto, el valor de la empresa será de 36 millones de dólares (21 millones de capital accionarlo más 15 millones de deuda); el cambio de la razón de deuda-capital accionario ha creado un valor nuevo de 6 millones de dólares. Pero algo debe de estar mal. Los activos netos de la empresa valían 30 millones de dólares originalmente porque generaban un ingreso de operación neto de 3 millones de dólares anuales. Nada ha ocurrido que vuelva a

[9] Franco Modigliani y Merton H. Miller, "The Cost of Capital, Corporation Finance and the Theory of Investment", 48 *Am. Econ. Rev.,* 261 (1958).

estos activos más productivos, de modo que ¿por qué pagaría un comprador ahora 36 millones de dólares por los activos?

No lo haría. La falacia del ejemplo hipotético es el supuesto de que la razón de precio-ganancias no se ve afectada por el monto de la deuda de la empresa. La adición de apalancamiento aumenta el riesgo del accionista porque la deuda es un costo fijo y cualquier cosa que incremente los costos fijos de una empresa la hace más vulnerable a una reducción en la demanda de su producto, por ejemplo, una contracción o desaceleración en el ciclo económico. Existe también un riesgo en el sentido ascendente —la empresa con un apalancamiento elevado se encuentra mejor que una firma provista de un apalancamiento inferior cuando la economía se recupera—; con todo, dado que la mayoría de los inversionistas sienten aversión por el riesgo, la razón de precio-ganancias de las empresas disminuirá a medida que aumenta su apalancamiento. En efecto, dado que el valor productivo de los activos de la empresa no se ve afectado por la forma en que decida la empresa arreglar los componentes de su estructura de capital, esperaríamos que la razón de precio-ganancias de la empresa bajara a 7.14, la razón a la que el valor de la empresa es igual que antes. No hay ninguna razón para que el comprador de una empresa pague una prima porque su estructura de capital contenga apalancamiento, aunque prefiera el rendimiento esperado elevado de una inversión de alto riesgo. Este inversionista puede crear su razón de deuda-capital accionario preferida comprando algo de las acciones con dinero que toma prestado. O puede emparejar su compra con la inversión en empresas muy apalancadas o de alto riesgo por alguna otra razón.

La pureza del análisis se ve afectada por el tratamiento fiscal diferencial de los intereses y las ganancias. El gobierno subsidia la adquisición de capital mediante préstamos permitiendo que las corporaciones deduzcan el costo del capital que se toma prestado, pero no el del capital accionario, de su ingreso gravable. Además, cuando la razón de deuda-capital accionario de una empresa es muy elevada, los riesgos para los tenedores de deuda y para los accionistas podrían volverse tan grandes que la empresa no pudiera atraer a ninguna de las dos clases de accionistas. Así pues, el monto del apalancamiento en su estructura de capital no es un asunto totalmente indiferente para el accionista (véase también antes XV.4); pero parece improbable que este factor pudiera explicar muchas fusiones de conglomerados.[10]

El análisis tiene implicaciones a la hora de evaluar el valor de una corporación para propósitos de cálculo de daños punitivos en casos en los que la

[10] Una explicación alternativa, basada en el mercado del control corporativo discutido en el capítulo anterior, destaca que las fusiones de conglomerados no se volvieron comunes antes del decenio de 1960, cuando se hicieron más severas las prohibiciones antimonopólicas contra las fusiones horizontales y verticales. Si no se incita a hacer compras horizontales y verticales de empresas subvaluadas, no se sigue de ahí que no se comprarán empresas en absoluto, sino antes bien que las adquirirán otras clases de compradores, tales como una empresa conglomerada.

mala conducta corporativa justifica imponer esa sanción. Resulta tentador suponer que el valor neto de la corporación es la medida adecuada del valor de la corporación. No obstante, el valor neto es un artefacto atribuible a una convención contable que maneja el capital de los accionistas de manera diferente a la de la deuda; si dos corporaciones tienen el mismo valor, pero una tiene una razón deuda-capital accionario elevada y la otra una razón deuda-capital accionario baja, la segunda corporación tendrá un valor neto más alto que la primera. La única justificación para imponer daños punitivos más onerosos a la segunda corporación que a la primera se fundaría en una creencia de que los accionistas pueden evitar la mala conducta con más presteza que los acreedores. Lo contrario podría ser cierto de una corporación que no tuviera accionistas conspicuos sino sólo unos pocos acreedores con un gran haber; cada uno de estos últimos podría ejercer más influencia sobre el comportamiento de la corporación que cualquiera de los accionistas.

XVI.3. ¿Por qué compran seguros las corporaciones?

Si los accionistas de una corporación pueden eliminar mediante la diversificación el componente de riesgo no sistemático de la corriente de ingresos esperados de la corporación, ¿por qué observamos que la mayoría de las corporaciones, grandes y pequeñas, públicas y cerradas, se aseguran contra las pérdidas resultantes de juicios de cuasidelitos, incendios, robo de los empleados y otras contingencias adversas? ¿Por qué los accionistas de una corporación pública no están adecuadamente asegurados sólo en virtud de que mantienen carteras diversificadas? Veamos algunas respuestas posibles:

1. Los administradores sienten aversión por el riesgo y tienen demasiado de su riqueza (incluido el capital humano) comprometido en la corporación para ser capaces de diversificar más allá de asegurar a la corporación. El seguro de la riqueza personal no es una alternativa viable, a causa del riesgo moral. El asegurado puede dejar de trabajar duro o, dado que no tiene ningún riesgo descendente, experimentar riesgos injustificados. De acuerdo con esta perspectiva, ¿se aviene la compra del seguro corporativo con los intereses de los accionistas, o se trata de otro ejemplo de costos de agencia? Es las dos cosas. La aversión de los administradores al riesgo abre una brecha entre sus intereses y los de los accionistas, pero reducir sus riesgos mediante la compra del seguro corporativo puede ser más barato que pagar a los administradores para que soporten ese riesgo.
2. El costo (administrativo) de la carga del seguro podría ser menor al costo de peso muerto esperado de la quiebra. En otras palabras, la quiebra contra la que el seguro protege, descontada por la probabilidad de

la quiebra si la corporación no está asegurada, costaría más en recursos reales consumidos (en honorarios legales, honorarios del administrador fiduciario y, sobre todo, en la pérdida de producción causada por la administración menos eficiente de una empresa que está operando bajo el control judicial) que la cantidad en la que la prima del seguro supera al pago esperado.

3. Las compañías de seguros son especialistas en la reducción de los riesgos mediante el asesoramiento sobre las precauciones de seguridad y también en el monitoreo de la defensa de ciertas reclamaciones. Si una corporación que no tiene un seguro contra sus responsabilidades es demandada según la doctrina del *respondeat superior* por un accidente causado por uno de sus empleados, tendrá que buscar un despacho de abogados para que la defienda de la acción cuasidelictiva, y tendrá que monitorear el desempeño de ese despacho de abogados. Todo esto será difícil si la corporación no tiene muchas reclamaciones de cuasidelitos en su contra. Como especialista en la supervisión de esa defensa, una compañía aseguradora de responsabilidades podrá arreglar la defensa contra tal reclamación a un costo menor y con mejores perspectivas de éxito que la corporación misma. Este análisis implica —y así lo observamos— que las corporaciones que afrontan reclamaciones cuasidelictivas repetitivas, tales como los ferrocarriles, con frecuencia se autoaseguran.
4. El seguro le permite a la compañía reducir la volatilidad de sus ganancias. Si los valores y los mercados de crédito no pueden distinguir fácilmente entre ganancias magras debidas a una mala administración y ganancias magras debidas a un choque exógeno del que no es responsable la administración, los administradores tendrán un incentivo para facilitar la obtención de ganancias a fin de evitar periodos en que éstas escaseen. En otras palabras, los administradores se aseguran contra la posibilidad de ser acusados por adversidades de las que en realidad no son culpables.

XVI.4. La selección de acciones y la tesis del mercado eficiente

Al concentrarnos en el diseño de la cartera y la estructura del capital, parecería que hemos perdido de vista la estrategia de inversión más clara —la compra de valores subvaluados, y la venta de valores sobrevaluados— y, lo que es lo mismo realmente, la cronología de los virajes del mercado. (Una creencia de que el mercado es "demasiado elevado" y está a punto de irse a pique es equivalente a la creencia de que [todas] las reservas están sobrevaluadas a su precio actual. Ocurre lo mismo con la creencia de que el mercado está vendido más de la cuenta y está a un paso de levantarse.) ¿Pero es ésta realmente una

estrategia eficiente? Por principio de cuentas, es algo cara. Se tienen los costos de investigación en que se incurre al seleccionar valores específicos para incluirlos en la cartera o excluirlos de ella, los costos de transacción en que se incurre al comprar y vender acciones de acuerdo con los resultados cambiantes del análisis de los valores, y, en fin, el sacrificio de la diversificación implicado por el mantenimiento de una cartera sustancialmente menor que la del mercado (la selección de acciones implica la selectividad, pero la diversificación requiere un enfoque incluyente). Para que todos estos costos se justifiquen, la selección de acciones deberá generar beneficios mayores, bajo la forma de un rendimiento esperado mayor, que el que pudiéramos esperar recibir de la cartera de mercado ajustada para imponer el mismo nivel de riesgo sistemático que la cartera creada por la selección de acciones.[11]

Podría parecer prácticamente evidente que un inversionista hábil, que realice una investigación cuidadosa de las condiciones y perspectivas de compañías particulares y del conjunto de la economía, ganará un rendimiento más elevado (siempre corrigiendo por las diferencias del riesgo sistemático) que el del inversionista que simplemente compra como el mercado, invirtiendo ciegamente en toda la lista del mercado de valores y no vendiendo jamás una acción cuando sus perspectivas empiezan a empeorar. Pero, dado que el valor de una acción es puramente una función de sus ganancias esperadas y, por lo tanto, depende en gran medida de los sucesos que ocurran en el futuro, determinar si una acción está sobrevaluada o subvaluada a su precio actual requiere saber lo que oculta el futuro, y muy pocas personas son hábiles para pronosticar el futuro. Aunque una acción podría estar subvaluada debido a alguna característica de la compañía (o de sus competidores, proveedores, clientes, el ambiente político, etc.) que existe ahora pero no se conoce generalmente o no se entiende correctamente, la información subyacente es del dominio público,[12] lo que quiere decir que está igualmente disponible para todos los analistas de valores. La única manera de ganar dinero con tal información consiste en interpretarla mejor que otros analistas. Éste no es un método promisorio para superar al mercado. Se requiere que el analista interprete la información disponible al público de un modo diferente al del promedio

[11] La calificación es vital: si el resultado de la selección de acciones es la creación de una cartera que tenga un beta mayor que 1, la cartera tendrá un rendimiento esperado mayor al de la cartera del mercado. Pero esto no se deberá a la habilidad del administrador de la cartera como seleccionador de acciones, sino al beta mayor de su cartera, que podría haberse alcanzado sin ninguna selección de acciones, simplemente apalancando la cartera del mercado hasta alcanzar el mismo beta.

[12] Si la información no es del dominio público, si es información confidencial, e importante, no podrá utilizarse legalmente, por lo menos por quienes tienen mayores probabilidades de obtenerla en privado, para comprar o vender acciones (véase antes XV.8). Presumiblemente, la ley tiene cierto efecto disuasivo… al menos hoy en día, dados los esfuerzos extraordinarios y, hasta la fecha, exitosos del Departamento de Justicia para detectar y perseguir a los violadores de la prohibición contra el tráfico de información confidencial.

de la opinión de la comunidad de analistas, y que sus interpretaciones diferentes sean correctas con una frecuencia sustancialmente mayor a la de sus interpretaciones incorrectas, con la finalidad sustancial de cubrir los costos de transacción implicados en la negociación de valores.

La dificultad de ganarle al mercado de forma consistente es incluso más grande de lo recién sugerido, debido a una cuestión que dejó entrever John Maynard Keynes. El corredor de la bolsa está menos interesado en el valor de una acción implicado por los "fundamentos" (qué tan bien administrada y financiada está la empresa, qué demanda tienen sus productos, etc.) que en lo que piensan sobre el valor que tiene el mercado los otros comerciantes u operadores de la bolsa, quienes a su vez pueden fundar su compra y sus decisiones de venta en lo que piensan que él piensa sobre el valor que posee el mercado, y así hasta el infinito en un círculo. Los fundamentos habrán de influir en las evaluaciones de los operadores de la bolsa y por ello en el precio de la acción, pero hay muchos otros factores que entran en juego y son en extremo difíciles de anticipar.

Confirmando estas observaciones teóricas, algunos estudios empíricos han descubierto que los fondos mutuos de inversión, a pesar de su empleo de analistas de valores y administradores de cartera a fin de superar al mercado, no lo logran.[13] No lo hacen mejor que la cartera de mercado a ciegas. Se ha sostenido que la comparación apropiada no es entre todos los fondos mutuos de inversión y el mercado, sino entre los fondos mutuos de inversión más exitosos y el mercado, pero los estudios sugieren que no hay fondos mutuos consistentemente exitosos. Algunos fondos mutuos disfrutan de corridas de éxito más cortas o más largas, pero generalmente el grado de éxito observado no es mayor que el que esperaríamos si la suerte, no la habilidad, fuera en efecto el único factor determinante del desempeño del fondo.

Los estudios apoyan una conclusión más fuerte aún: cuando se toman en cuenta los costos del corretaje y los honorarios de los administradores, el fondo mutuo o el fondo de fideicomiso común tienen un rendimiento neto menor que el de un índice de mercado de base amplia tal como el S&P 500.[14] Esta comparación fue blanco de burlas durante largo tiempo alegando que el S&P

[13] Los primeros estudios se resumen en Lorie y Hamilton, XVI.1, antes nota 4, en el cap. 4 de su libro. Véanse también Burton G. Malkiel, "Reflections on the Efficient Market Hypothesis: 30 Years Later", 40 *Fin. Rev.,* 1 (2005); R. A. Brealey, *An Introduction to Risk and Return from Common Stocks*, cap. 3 (2ª ed., 1983); Richard A. Ippolito, "Efficiency with Costly Information; A Study of Mutual Fund Performance, 1965-1984", 104 *Q. J. Econ*., 1 (1989). La investigación empírica se ha concentrado en los fondos mutuos de inversión porque en los términos de la ley federal están obligados a reportar con detalle su desempeño; pero todo indica que los fondos de fideicomiso comunes, los fondos de pensiones y otros inversionistas institucionales también fracasan sistemáticamente en su intento de superar a la cartera del mercado.

[14] Véanse, por ejemplo, Burton B. Malkiel, "Passive Investment Strategies and Efficient Markets", 9 *Europ. Fin. Mgmt.*, 1, 3, 10 (2003); Michael C. Jensen, "Risk, the Pricing of Capital Assets, and the Evaluation of Investment Portfolios", 42 *J. Bus.,* 167 (1969), y referencias en Lynn A.

500 es un fondo hipotético y, por lo tanto, no tiene costos administrativos. Ahora que están —y en realidad han estado a lo largo de décadas— en operación algunos fondos reales basados en el mercado (véase después XVI.6), es posible evaluar esta crítica… y rechazarla. Los costos administrativos de un fondo de mercado resultan ser tan bajos (en una cartera de 500 millones de dólares, quizá 10% de los costos de la administración convencional) que el rendimiento esperado de un fondo de mercado apropiadamente construido es sólo trivialmente diferente del rendimiento del S&P 500.

XVI.5. El desafío de las finanzas del comportamiento

Los varios puntos planteados en este capítulo en lo tocante a la forma de operar de los mercados de valores se apoyan en la noción de que éstos funcionan como si los operadores de la bolsa fuesen racionales (lo que por supuesto no significa que sean omniscientes: la información es costosa y, por lo tanto, el término "ignorancia racional" no es un oxímoron). El "como si" del supuesto debe ser enfatizado. Podría ser que muchos operadores de la bolsa fueran irracionales ("comercian con base en el ruido" en vez de comerciar con fundamento en la información) y, de todos modos, si sus desviaciones de la racionalidad fuesen aleatorias, el precio promedio de un valor no sería afectado. Y, aun cuando esas desviaciones fuesen sistemáticas, el efecto sobre el precio promedio sería ligero si los arbitrajistas se mantuviesen alerta en pos de oportunidades de regatear. Los arbitrajistas son especuladores que buscan situaciones en las que dos activos idénticos se venden a diferente precio. Un arbitrajista podría, por ejemplo, explotar la presencia de una discriminación de los precios (cuando el mismo producto se vende a precios distintos a compradores que tienen diferentes elasticidades de demanda del mismo) comprando el producto de los compradores a quienes se les cobra un precio más bajo debido a que su demanda es más elástica y vendiéndoles a los compradores a quienes se les cobra un precio más alto porque su demanda es menos elástica.

Supongamos que, a causa de un apego irracional por el valor *X* por encima del muy similar valor *Y*, el precio de *X* se eleva respecto al de *Y* aun cuando el rendimiento esperado de los dos valores es el mismo. Al vender *X* en corto y comprar *Y*, el arbitrajista realiza una ganancia más o menos garantizada puesto que, si los dos valores son en efecto sustitutos cercanos, un aumento continuo en el precio de mercado de *X* probablemente irá acompañado de un aumento en el valor de mercado de *Y*, de manera que el arbitrajista recuperará en ganancias a partir de *Y* lo que pierda si, contrario a sus expectativas, el precio de x no baja. Si en efecto baja, su venta en corto será rentable, y es

Stout, "Are Stock Markets Costly Casinos? Disagreement, Market Failure, and Securities Regulation", 81 *Va. L. Rev.,* 611, 623 n. 30 (1995).

improbable que incurra en una pérdida proporcional con *Y*, dado que *Y* estaba subvaluado en relación con *X* cuando se vendió por menos (toda vez que los valores son tan similares). La existencia de un sustituto cercano de *X*, a saber, *Y*, es lo que le permite al arbitrajista tener una cobertura, pues sin dicho sustituto su venta en corto sería muy arriesgada, ya que no puede estar seguro de que ese valor se encuentre sobrevaluado; dicho con más precisión, no puede tener la certeza de que el mercado "se despertará" y se percatará de que está sobrevaluado.

Una literatura creciente sobre las "finanzas del comportamiento" —esto es, una literatura que encuentra su mayor sostén en las profundas ideas e intuiciones de la psicología cognitiva vista como lo que se contrapone al modelo de comportamiento racional empleado en el análisis económico convencional y es, así, prima hermana de la economía del comportamiento— pone en duda, provista de una base empírica no poco significativa, los supuestos clave de la teoría de los mercados eficientes. Entre los supuestos cuestionados se halla la eficacia del arbitraje en la neutralización del comportamiento irracional en materia de inversiones (un comportamiento ilustrado por las diferencias arbitrarias en el precio de los valores que son mutuamente sustitutos cercanos).[15] "Los inversionistas siguen el consejo de los gurús financieros, no logran efectuar una diversificación, comercian activamente con acciones y remueven sus carteras, venden acciones ganadoras y se aferran a acciones perdedoras provocando un incremento en sus responsabilidades fiscales, compran y venden fondos mutuos administrados de forma activa y a precios elevados, siguen de cerca las pautas observadas en los precios de los valores y otros modelos populares."[16] Por supuesto, todo esto es bien sabido desde antaño. Pero la literatura sobre las finanzas del comportamiento descubre que tales comportamientos son sistemáticos antes que aleatorios. Los inversionistas se muestran más renuentes a vender los valores perdedores que los ganadores ("aversión a la pérdida"). Abrigan la expectativa exagerada de que las tendencias habrán de proseguir, y esta expectativa amplifica las oscilaciones del mercado de valores al alentar a hacer la compra cuando los precios de los valores están al alza y a venderlos cuando están a la baja. Asimismo, los inversionistas demandan una prima de riesgo más alta de lo que la aversión al riesgo justificaría (dada la posibilidad de reducir el riesgo mediante la diversificación) por poseer valores en relación con los bonos, debido a que los valores presentan un mayor riesgo en sentido descendente y la aversión al riesgo

[15] Para un panorama lúcido aportado por una figura cimera en el campo, véase Andrei Shleifer, *Inefficient Markets: An Introduction to Behavioral Finance* (2000). Véase asimismo Nicholas Barberis & Richard Thaler, "A Survey of Behavioral Finance", en *Handbook of the Economics of Finance,* 1054 (G. M. Constantinides, M. Harris & R. Stulz [comps.], 2003). Para una defensa vigorosa de la teoría de los mercados eficientes, véase Burton G. Malkiel, "The Efficient Market Hypothesis and Its Critics", *J. Econ. Perspectives*, 59, invierno de 2003.

[16] Shleifer, nota 15 antes, en la 10.

implica que el riesgo descendente tiene un peso mayor en la decisión del inversionista que el riesgo en sentido ascendente. Y, puesto que las personas, incluidos muchos inversionistas, se topan con dificultades al manejar probabilidades y, por consiguiente, tienden a no entender que los plazos o "corridas" guardan consistencia con el azar (lo cual se relaciona con la inclinación de las personas a esperar que las tendencias prosigan de manera indefinida), perciben pautas donde no las hay y, por ende, le otorgan al desempeño de los valores a corto plazo —así como al desempeño a corto plazo de los administradores de dinero— un peso mayor del que merecerían recibir.

Los administradores de dinero o corredores de divisas que tienen un perfil profesional —o al menos muchos de ellos (pues los profesionales no son inmunes a defectos cognitivos)— saben mejor del asunto. No obstante, si están compitiendo por los dólares de inversión procedentes de inversionistas poco sofisticados, deben satisfacerlos en lo concerniente a sus predilecciones irracionales. (Esta cuestión se relaciona con el punto anterior acerca del comercio efectuado con base en lo que el comerciante cree que piensan los demás comerciantes.) No se puede depender del arbitraje para que estas irracionalidades queden eliminadas porque el arbitrajista a menudo se ve imposibilitado para procurarse una cobertura, y cuando no lo logra el arbitraje llega a ser en extremo riesgoso. No puede cubrirse si no hay buenos valores que sean sustitutos de los que él juzga sobrevaluados o subvaluados. E, incluso si los hay, puede ocurrir que el mercado no despierte a tiempo para que la cobertura funcione. En el ejemplo que ponemos, si el precio de *X* se sigue elevando después de que el arbitrajista lo ha vendido en corto y ha comprado *Y*, pero *Y* no se eleva al unísono, puede quedar expuesto a pérdidas colosales antes de que los dos precios de valores converjan finalmente. Dado que el arbitraje es incompleto, las irracionalidades de los inversionistas persisten y desencadenan desviaciones sistemáticas entre el precio de los valores y el valor fundamental. La conducta irracional del inversionista también es promovida por los fondos mutuos, los corredores de la bolsa y otros profesionales de los valores que perciben que al fomentar dicho comportamiento obtienen oportunidades de ganancia. Hasta qué grado ocurren estas desviaciones es un asunto que ha suscitado controversias.[17] Y, tal como suele suceder con los hondos hallazgos de la psicología cognitiva, las implicaciones en cuestión de políticas resultan poco claras. Lo que es innegable es que, entre más incautos son los inversionistas, más importancia cobra tener a la mano remedios efectivos contra el fraude en valores: supongamos, por ejemplo, que si un fondo mutuo se promocionara anunciando que el hecho de que durante dos años haya observado de manera sostenida un desempeño por encima del promedio prueba que su desempeño en el tercer año será también superior al promedio, la ley se ne-

[17] Burton G. Malkiel, "Searching for Rational Investors: Explaining the Lowenstein Paradox", 30 *J. Corp. L.,* 567 (2005); G. William Schwert, "Anomalies and Market Efficiency", en *Handbook of the Economics of Finance*, antes nota 15, en 937-972.

garía a concederles a los inversionistas crédulos un remedio por la razón de que nadie podría ser tan estúpido para creer tal cosa. Con todo, ir mucho más allá de esta postura y, por dar un caso, prohibirle a la gente comprar valores sin antes haber pasado una prueba que demuestre que piensan con claridad parece algo difícilmente viable. Las limitaciones cognitivas que están enraizadas en lo más profundo de la mentalidad humana desafían a todo remedio legal. Aun así, es importante que la ley no dé por sentado que las personas no son más racionales de lo que en verdad son.

XVI.6. UN RETO MAYOR

En su versión más deslucida, que es todo lo que se requiere para mostrar que la selección de acciones constituye probablemente una estrategia inferior a la de comprar y mantener una cartera diversificada, la teoría de los mercados eficientes afirma simplemente que el mercado o la bolsa de valores es difícil de ganar porque hay muchos operadores de la bolsa esforzándose en lograrlo y todos tienen acceso a la misma información. En su versión más sólida, la teoría afirma que los precios actuales de los valores son la mejor estimación de las ganancias futuras de las compañías cuyas acciones están negociándose.[18] La teoría de las finanzas del comportamiento podría convencernos de poner "una estimación bastante buena" en lugar de "la mejor estimación", pero una crítica más radical de la versión sólida de la teoría de los mercados eficientes es que con frecuencia los precios actuales de los valores pueden hacer predicciones seriamente erróneas sobre las ganancias corporativas, y ello por razones que no se fundan en el planteamiento según el cual muchos comerciantes u operadores de bolsa son irracionales.

Las personas que negocian con valores, colocadas aparte de las personas que los compran y mantienen consigo por un largo plazo, no están interesadas principalmente en las ganancias corporativas que obtengan en el futuro las compañías con cuyos valores están negociando; su interés primordial estriba en saber si otros negociantes u operadores creen en la probabilidad de que haya un alza o una baja en el valor de las acciones; y esos otros negociantes u operadores están interesados igualmente en lo que otros negociantes u operadores más creen al respecto. Un negociante que piensa que muchos otros negociantes juzgan que un valor está subvaluado tiene buenas razones para comprarlo, sin importar lo que piense sobre cuáles serán probablemente las ganancias futuras de la compañía. De ahí el "comercio movido por el ímpetu": comprar cuando los demás están comprando, vender cuando los demás

[18] La mejor, pero no la perfecta: "Debido a que la información es costosa, los precios no pueden reflejar de modo perfecto la información que está disponible, ya que, si lo hicieran, los que gastan recursos para obtenerla no se verían compensados". Sanford J. Grossman y Joseph E. Stiglitz, "On the Impossibility of Informationally Efficient Markets", 70 *Am. Econ. Rev.*, 393, 405 (1980).

están vendiendo. Esta tendencia (que debería uno distinguir de la tendencia antes analizada a exagerar el carácter persistente de las tendencias) ha sido motivo de mofas y calificada como "comportamiento de la manada", conducta que puede dar la impresión de ser irracional pero que no lo es, y no solamente entre los (otros) animales. (Si eres un antílope y adviertes que los otros antílopes de súbito emprenden una ruidosa estampida, estás bien avisado al determinar sumarte a ellos porque pueden estar huyendo de un león o de otro depredador.) El comercio movido por el ímpetu constituye un comportamiento de manada de carácter racional cuando tiene como fundamento una conjetura racional sobre el comportamiento de otros negociantes u operadores, aun cuando también puede reflejar la tendencia humana a ver pautas donde no hay ninguna de ellas (posiblemente debido a que la detección de pautas es un rasgo humano producto de la evaluación y resulta de inmenso valor en la mayoría de las situaciones). Empero, el comercio movido por el ímpetu es también peligroso para la economía en su conjunto: puede dar lugar a burbujas en el precio de los activos. Acontecimiento causante de desequilibrio, una burbuja involucra un incremento pronunciado en el precio que persista durante un tiempo considerable, no puede hallar explicación en los principios fundamentales y, después de alcanzar un pico, da rápidamente paso a una caída abismal en el precio. Véase antes XVI.4. El estallido de la burbuja puede connotar una crisis económica general, como aprendimos en 2008.

He aquí un ejemplo del comportamiento de manada de signo racional: Supongamos que, de acuerdo con el consenso entre los analistas de valores, la acción de una compañía está subvaluada. Usted es un analista de valores y no está de acuerdo; piensa que está sobrevaluada. ¿Va a recomendar a sus clientes que vendan el valor? Probablemente no, a menos que o bien tenga la suma seguridad de que tiene razón o sea sumamente osado. (El temperamento difiere entre las personas, y el temperamento influye sobre las acciones… en particular bajo una sombra de incertidumbre, en la que los datos "objetivos" no pueden servir de guía fiable a las decisiones.) Puesto que, si no está en lo cierto, lo colocarán en un sitio lejos de los demás y será el blanco de las críticas. Si a pesar de sus dudas aconseja comprar la acción y luego resulta estar equivocado, es usted parte de la manada: donde todos se hallan en el error, no hay nadie a quien culpar. Así pues (dependiendo de su estimación de las probabilidades), la decisión que maximiza su ingreso esperado puede estar en consonancia con la manada por más que manifieste escepticismo ante la sabiduría de la manada. Supóngase que la probabilidad de que su recomendación sea correcta es p, de modo que la probabilidad de que sea incorrecta es $1 - p$; y supóngase que el beneficio para usted respecto al ingreso futuro derivado de una reputación en aumento es x y la pérdida para usted respecto al ingreso futuro derivada de una reputación en mengua es y; entonces recomendará la compra si $px > (1 - p)y$, y recomendará la venta si $px < (1 - p)y$.

Se ha observado que los precios de los valores tienden a sumirse al final de cada trimestre —una pauta que no puede explicarse por medio de cambios en las ganancias corporativas predichas o, para el caso, por el comercio movido por el ímpetu—. Existe una explicación irracional, fundada en los costos de agencia, aunque de nuevo se trata de una explicación que arroja dudas sobre la versión sólida de la teoría de los mercados eficientes. El final del periodo trimestral (o el final del año, que es el término de un trimestre) es el momento en que usualmente el desempeño de un administrador de carteras se somete a una revisión a cargo de sus clientes o superiores. Si su cartera contiene una cantidad de acciones cuyo valor ha disminuido a lo largo del trimestre, tendrá que dar algunas explicaciones; y, en lugar de tener que hacer eso, puede decidirse a vender esos valores perdedores antes de la revisión trimestral. La incertidumbre desempeña un papel en este ámbito porque, si los administradores de carteras fuesen buenos seleccionadores de acciones, un administrador no se vería impelido a tratar de explicar con suficiencia el porqué de tantos perdedores. Comoquiera que sea, según sabemos, hay algunos buenos seleccionadores de acciones.[19]

El desafío —esbozado en esta sección— que se plantea a la versión sólida de la teoría de los mercados eficientes desempeñará un papel estelar cuando, más adelante en este capítulo, nos centremos en el colapso financiero de 2008.

XVI.7. Implicaciones de la teoría financiera para el monopolio regulado

La teoría del mercado de capital puede ayudar a resolver el problema de la limitación de las ganancias de los monopolistas regulados. El problema se complica por la dificultad de determinar el costo del capital accionario, una dificultad generada a su vez por la ausencia de un precio directamente observable. El enfoque regulador actual es en gran medida circular pues basa la tasa de rendimiento permitida en las tasas de rendimiento de otras compañías reguladas. En cambio, si los reguladores empezaran por medir el beta de las acciones de la empresa regulada, podrían fijar un nivel de ganancias permitido tal que el rendimiento esperado para los inversionistas de la empresa fuese igual al de los inversionistas de las empresas no reguladas del mismo beta. Éste sería el costo verdadero, para la empresa regulada, de atraer capital accionario sin diluir el valor de las acciones propiedad de los accionistas existentes.

La teoría del mercado de capital podría ser útil también para desviar a las dependencias reguladoras (y sus críticos) de lo que son en gran medida

[19] Para otras anomalías en la teoría de los mercados eficientes, véase Yvan Lengwiler, *Microfoundations of Financial Economics: An Introduction to General Equilibrium Asset Pricing,* caps. 7-9 (2004); Jean-Pierre Danthine y John B. Donaldson, *Intermediate Financial Theory,* 140-141 (2ª ed., 2005).

problemas espurios, como el de la razón óptima de deuda-capital accionario. En el caso de una empresa regulada con una razón deuda-capital accionario baja, se sostiene a menudo que si sólo la empresa aumentara la razón su costo de capital bajaría —porque la tasa de interés es menor al rendimiento de las acciones comunes— y sus precios podrían ser disminuidos. Pero esto equivale a sostener que las empresas reguladas debieran estar muy apalancadas. Dado que la deuda aumenta la volatilidad del capital accionario, una empresa que aumentara la proporción de deuda en su estructura de capital experimentaría un aumento en el costo de su capital accionario. No hay ninguna razón para pensar que el costo global del capital sería menor —o, para el caso, más elevado— con una razón mayor de deuda-capital accionario (véase antes XVI.2).

Esta conclusión puede parecer vulnerable a las críticas de los teóricos de las finanzas del comportamiento. La idea de que la estructura del capital (la razón deuda-capital accionario) es una cuestión indiferente más allá de las consideraciones fiscales da por sentado que las acciones de las diferentes compañías son perfectamente sustituibles. Si ese supuesto se atenúa[20] y algunos inversionistas están ávidos, digamos, de poseer acciones en la compañía X, un decreto regulatorio que estipulara que X debe incrementar su razón deuda-capital accionario podría reducir el valor de la compañía en lugar de, como se suponía más arriba, dejarlo invariable. Si en cambio la compañía les resultara atractiva a los inversionistas en bonos pero no a los inversionistas en acciones, un aumento en la razón deuda-capital accionario podría reducir el costo de capital general de ésta. Empero, es de presumir que la compañía estaría al tanto de ello y escogería su razón deuda-capital accionario en consonancia con la situación, sin necesidad de que la incitara a hacerlo la dependencia regulatoria.

XVI.8. UNA REVISIÓN DEL ESOP: EL PROBLEMA DE LA SUBDIVERSIFICACIÓN

En el capítulo previo señalamos que el ESOP o Plan de Propiedad de Acciones de los Empleados supera un problema serio que ha asolado a las cooperativas de trabajadores —el horizonte trunco de los trabajadores—, pues el empleado que participa en ese proyecto tiene un interés como propietario, dado que sus acciones reflejan el valor presente de las ganancias futuras de la empresa, incluidas las ganancias habidas una vez que él renuncia o se retira. No obstante, hay un problema de subdiversificación si las acciones del ESOP en poder del empleado constituyen un porcentaje sustancial de su riqueza.[21] A causa de

[20] Véase Andrei Shleifer, *Inefficient Markets: An Introduction to Behavioral Finance,* 185 (2000).

[21] Éste es un problema general con los empleados que poseen valores en el ámbito de su empleador, como resulta común en los planes de retiro de 401 (mil) dólares. Véase Lisa Meulbroek, "Company Stock in Pension Plans: How Costly Is It?", 48 *J. Law & Econ.,* 443 (2005).

este problema, la posición del administrador fiduciario que maneja el ESOP se hace incómoda. Por un lado, no puede tener ningún deber de diversificación como tal, porque esto sería inconsistente con el concepto del ESOP. Por otro lado, ERISA le impone un deber de prudencia. El choque entre estos principios se ilustra con el destino que le tocó a UAL, casa matriz de United Airlines. Entre el inicio de 2001 y septiembre de 2002, cuando UAL estaba al borde de la quiebra (la declaró en diciembre de 2002), el precio de sus acciones cayó de 45 a poco más de 2 dólares. No fue hasta septiembre de 2002 que el administrador fiduciario empezó a vender acciones de UAL procedentes del ESOP, y para entonces ya era muy tarde para evitar una pérdida desastrosa de riqueza a los participantes en el ESOP. No sólo eran sus acciones en el ESOP el mayor activo financiero de muchos de sus participantes, sino que, en tanto empleados de United Airlines, su caudal no financiero —sobre todo ganancias esperadas y prestaciones accesorias— estaba positivamente correlacionado con el ingreso de United y, de este modo, se derrumbó con la bancarrota de la compañía, exacerbando así las repercusiones de la subdiversificación.

Los participantes demandaron al administrador fiduciario con el argumento de que debería haber descargado las acciones de UAL tan pronto como los ataques terroristas del 11 de septiembre de 2001 hicieron patente que las líneas aéreas se encontraban en un serio peligro financiero.[22] Era un mal argumento, y los demandantes perdieron el caso. Un mes después de los ataques del 11/9, el mercado aún valuaba las reservas de UAL a 15 dólares la acción. Si el mercado hubiera sospechado que UAL estaría en quiebra para finales de 2002, el precio de las acciones de UAL no habría estado tan elevado en octubre de 2001, con la implicación de que a la compañía le correspondía una capitalización de mercado de más de 800 millones de dólares. Un administrador fiduciario no comete la imprudencia de suponer que un mercado de valores de gran envergadura (United se incluía en las negociaciones de la Bolsa de Valores de Nueva York) proporciona la mejor estimación disponible del valor de las acciones negociadas en él.[23] A *cualquier* precio en que una venta se verifica, el comprador piensa que está consiguiendo una ganga. Es verdad que algunos inversionistas logran anotarse un sólido triunfo en el mercado, pero sólo en casos contados se trata de administradores fiduciarios; si el responsable de un fideicomiso cree que es capaz de predecir el futuro de UAL con más precisión que el mismo mercado, está acarreando la soberbia sobre la entidad a su nombre.

No obstante, la caída en el precio de mercado de UAL estaba aumentando el riesgo arrostrado por los dueños de sus reservas, los participantes del ESOP. Como lo sabemos, entre mayor es la razón de la deuda de intereses fijos al capital accionario, más riesgosa es la posición de los tenedores de dicho ca-

[22] Véase In re UAL Corp., 443 F.3d, 565 (7th Cir., 2006).

[23] Burton G. Malkiel, "Reflections on the Efficient Market Hypothesis: 30 Years Later", 40 *Fin. Rev.,* 1 (2005).

pital (los accionistas comunes). Conforme el valor de UAL —según los cálculos hechos en el mercado de valores— se iba desplomando, la razón deuda-capital accionario de la compañía se hizo estratosférica porque su deuda no decrecía, y de este modo el peligro enfrentado por los participantes en el ESOP aumentó más allá de lo esperado al momento de la creación del plan. En algún punto del deslizamiento, por lo tanto, el deber de la prudencia podría haberse antepuesto a la consideración de que el administrador fiduciario de un ESOP no tiene por qué dictar la diversificación, pero es de dudar que la corte fuera capaz de determinar cuál era el punto.

XVI.9. La Ley de la Inversión en Fideicomisos y los fondos de índices

La ley de fideicomisos requiere que un administrador fiduciario, al administrar fondos de fideicomisos, actúe prudente y cautamente con el propósito primordial de preservar el principal del fideicomiso. Se supone que los beneficiarios del fideicomiso sienten gran aversión por el riesgo y, por lo tanto, prefieren recibir un rendimiento esperado menor a cambio de asumir riesgos menores. El supuesto es razonable en lo tocante a muchos fideicomisos, aunque no a todos ellos. Si los beneficiarios son una viuda e hijos menores de edad que no tienen otro ingreso, y sí una capacidad limitada para ganarlo, y a resultas de inversiones riesgosas del principal baja marcadamente el ingreso del fideicomiso, la posición de los beneficiarios puede volverse en extremo incómoda. Las inversiones riesgosas son más atractivas para quienes tienen salarios regulares u otras fuentes de ingreso básico estables que protegen su nivel de vida si las inversiones fracasan… pero por supuesto algunos beneficiarios de fideicomisos se encuentran sujetos a esta eventualidad. El creador de un fideicomiso que desea que el administrador fiduciario haga inversiones riesgosas puede establecerlo así mediante un lenguaje apropiado en la escritura constitutiva del fideicomiso. La regla del hombre prudente desempeña la función, ahora ya familiar, de reducir los costos de transacción implicando en toda escritura una cláusula (a menos que se diga lo contrario) que la mayoría de las partes incorporarían de otro modo mediante un lenguaje expreso.

Dicha regla, y con ella la ley de inversiones de fideicomisos en general, han adquirido creciente importancia desde la promulgación de la Ley de Seguridad de los Ingresos de Jubilación del Empleado (ERISA) en 1976 (véase antes XII.11). ERISA impone la regla del hombre prudente sobre los administradores de fondos de pensión y prohíbe que esta regla se pase por alto.

Si bien el principio general subyacente en la regla del hombre prudente tiene sentido económico, algunas de las reglas de ejecución tradicionales (hoy por hoy, no obstante, felizmente abandonadas), en particular la que aplica el estándar a las inversiones individuales antes que al total de la cartera,

no lo tienen. De acuerdo con este enfoque, por bien que se desempeñe la cartera, el administrador fiduciario se halla en riesgo de ser considerado responsable del mal desempeño de una de las inversiones de la cartera si no verifica la conveniencia de esa inversión antes de hacerla. Este enfoque tiene tres consecuencias negativas:

1. Los administradores fiduciarios se ven inducidos a gastar tiempo y dinero en la investigación de las perspectivas de valores individuales aunque los costos de la búsqueda y la ejecución implicados en la identificación y la adquisición de valores subvaluados para comprarlos, y en la revisión continua de nuestra cartera en búsqueda de valores sobrevaluados para venderlos, superan casi siempre a los beneficios, de acuerdo con el análisis económico de la selección de acciones.
2. Los administradores fiduciarios se abstienen de invertir en valores enteramente buenos sólo porque la compañía que los emitió tiene malas perspectivas de ganancias. En la teoría económica no hay ninguna presunción de que los precios de las acciones de las empresas declinantes o aun en quiebra estén típicamente sobrevaluados. Se hará que tales precios bajen hasta el punto en que las ganancias esperadas de la compañía (cualesquiera que sean o puedan ser) generen para el inversionista un rendimiento normal de su inversión en valores con la volatilidad de estas acciones (beta), la que podría ser baja.[24] Aun así, podría considerarse según la regla tradicional del hombre prudente que un administrador fiduciario que comprara conscientemente las acciones de una empresa quebrada actuó imprudentemente.
3. La aplicación de la regla del hombre prudente a cada inversión sucesiva induce a los administradores fiduciarios a mantener carteras insuficientemente diversificadas. Si el administrador fiduciario debe investigar cada acción que compre y vigilarla estrechamente tras de la compra, será limitado el número de acciones diferentes que puede mantener en la cartera. Inconsistentemente, la ley, aun cuando se aplica según una base de inversiones individuales, requiere que los administradores fiduciarios diversifiquen sus carteras; los casos no especifican el grado de la diversificación que se requiere. Esto colocaba a los administradores fiduciarios en la misma posición poco cómoda en la que se halla un fiduciario de la ESOP aún hoy en día, cuando, como estamos a punto de ver, la regla del hombre prudente ha sido modernizada.

Una manifestación de la preocupación tradicional de la ley por el cuidado en la selección de inversiones individuales es el concepto, incorporado en

[24] La empresa podría tener un valor positivo aunque sus ganancias actuales sean cero o negativas. Sus activos podrían tener un valor residual o valor de salvamento, o la empresa podría tener la perspectiva de ganancias en el futuro.

muchas leyes reglamentarias y en las decisiones judiciales del Estado, de que categorías enteras de inversiones son ilegales para un administrador fiduciario, a menos que se lo permita expresamente la escritura constitutiva del fideicomiso. Durante largo tiempo no se permitió que los administradores fiduciarios invirtieran en acciones comunes. Esto los llevó naturalmente a invertir fuertemente en bonos. Lo que se dejó de lado fue el hecho de que un bono de largo plazo expone al tenedor a un riesgo que evitaría si invirtiera en acciones (o en un instrumento de deuda a corto plazo): el riesgo de que la inflación aumente. Un bono es una protección enteramente buena contra la tasa de inflación *esperada*, porque la tasa de interés fijada en el bono incluirá la tasa inflacionaria como un componente. Pero si la tasa inflacionaria aumentara inesperadamente durante el periodo del bono, la disminución resultante del valor real del bono (descontada la inflación) será soportada enteramente por el tenedor. Un inversionista que sienta aversión por el riesgo no deseará soportar tal riesgo y no tiene que hacerlo.

Teniendo en mente la repetida solicitud de que el administrador fiduciario monitoree cada uno de los valores, la ley de algunos estados de la Unión Americana limita la autoridad de éstos para comprar acciones de fondos mutuos. Empero, cuando los activos de un fideicomiso son pequeños, podría resultar imposible alcanzar una diversificación razonable mediante procedimientos distintos de la compra de acciones de un fondo mutuo. La base de la regla era la premisa falsa de que al comprar acciones de fondos mutuos abdica el administrador fiduciario su responsabilidad fundamental de seleccionar las inversiones para el fideicomiso a favor de los administradores del fondo mutuo.

Dado que la ley se encuentra tan desfasada respecto de una comprensión económica, uno esperaría que los redactores de las escrituras constitutivas del fideicomiso hubieran incluido de ordinario un lenguaje que eliminara las limitaciones que había impuesto la ley de los fideicomisos a las reglas de incumplimiento u omisión en cuanto a inversiones. Y, en efecto, incluso antes de las recientes reformas legales, la mayoría de las escrituras constitutivas de fideicomisos eliminaban muchas de las detalladas limitaciones de la ley de fideicomisos y otorgaban al administrador fiduciario una discreción amplia, lo que explica la escasez de los litigios reportados que impliquen las obligaciones de inversión de los administradores fiduciarios.[25] Pero finalmente la ley se puso al corriente con la teoría moderna de las finanzas. Hoy en día todos los estados han transformado la regla del "hombre prudente" en la regla del

[25] Entre más barata resulte la contratación al margen de una regla del derecho común ineficiente, más se tardará el cambio de la regla mediante la litigación; es más, si la contratación al margen es suficientemente barata, no habrá ninguna litigación sobre la regla. Jeffrey N. Gordon, "The Puzzling Persistence of the Constrained Prudent Man Rule", 62 *N. Y. U. L. Rev.*, 52, 75-88 (1987).

"inversionista prudente", que exige que los administradores fiduciarios ajusten sus prácticas de inversión a las enseñanzas de esa teoría.[26]

Los administradores fiduciarios ahora colocan en fondos indexados porciones de los fondos bajo su cuidado. Un fondo indexado típico compra y mantiene una cartera de 200 a 500 valores destinados a igualar el desempeño de la Bolsa de Valores de Nueva York (y quizás algún promedio ponderado de mercados de valores nacionales y extranjeros) sin realizar ningún análisis de los valores y negociando sólo en la medida en que sea necesario para mantener la diversificación, manejar las amortizaciones e invertir el efectivo de sus accionistas. Es un epítome de la inversión pasiva. Pero esto provoca la siguiente interrogante: ¿qué ocurriría si todos los inversionistas adoptaran la estrategia pasiva implicada por el concepto? El mercado cesaría de ser eficiente entonces. Pues, en la medida en que es eficiente, lo es sólo porque muchos miles de inversionistas y sus asesores están en busca de información que revele cuáles valores están sobrevaluados o subvaluados y reaccionan con prontitud ante las noticias obtenidas. No hay aquí motivo para preocuparse; mucho antes de que cada inversionista hubiera adoptado la estrategia pasiva, algunos inversionistas la desechaban para aprovecharse de las crecientes oportunidades que había de obtener ganancias positivas a base del análisis de valores y la negociación activa. La observación de otros mercados, por ejemplo, el mercado de la vivienda, donde las transacciones son relativamente infrecuentes y los productos negociados son heterogéneos (no hay dos casas tan iguales como dos acciones de la misma clase de acciones de la reserva de la misma compañía), sugiere que la bolsa de valores seguiría siendo eficiente aunque la mayoría de los inversionistas fuesen pasivos.

Una segunda interrogante acerca de los fondos indexados se refiere a sus implicaciones potenciales para la gobernanza corporativa en general. Si la mayoría de los accionistas de una corporación fuesen inversionistas pasivos que nunca vendieran sus acciones cuando alguien ofrecía un precio elevado por ellas —en lugar de ello suponiendo que sus tenencias estaban siempre correctamente valuadas al precio corriente (y ahora mayor)—, ¿cómo sería posible adquirir una corporación por la vía de la oferta de compra o aun mediante una lucha de poderes? Si los fondos de mercado devinieran los principales inversionistas corporativos y simplemente hicieran caso omiso del material sustituto —un corolario del enfoque pasivo, cabe sostener—, el control de los accionistas sobre los administradores se debilitaría. Pero, por otro lado, el debilitamiento no sería de gran escala hasta que una fracción numerosa de valores corporativos se encontrara bajo la administración pasiva, y en ese punto reaparecerían las ganancias derivadas del estudio del material sustituto.

[26] Véase Restatement (Third) of Trusts, 227 (1992); Max M. Schanzenbach y Robert H. Sitkoff, "Did Reform of Trust Investment Laws Change Trust Portfolio Alternatives?", 50 *J. Law & Econ.*, 681 (2007). La respuesta de los autores a la interrogante que aporta su título, dada tras un cuidadoso análisis empírico, es "sí".

La crisis económica de 2008 (y continua) ha suscitado preguntas sobre el viraje de los bonos a las acciones como los componentes fundamentales de la cartera de inversiones de un fideicomiso.[27] Ahora es posible percatarse con mayor agudeza del punto formulado al comienzo del capítulo de que no todos los riesgos son diversificables; en particular, el riesgo que es común al mercado de valores en su integridad (beta) no puede ser diversificable. No sólo los precios de las acciones han sido extremadamente volátiles desde la burbuja de las empresas digitales (.com) de fines de la década de 1990, sino que el valor de los capitales sociales se ha estancado en conjunto desde el desplome subsiguiente. Y ahora sabemos que los precios de las acciones no pueden ser tomados como simples estimaciones de las ganancias corporativas futuras, pues son vulnerables a las burbujas que pueden conducirlos a alturas insospechadas seguidas súbitamente por declives drásticos. La impredecibilidad y la inestabilidad del mercado de valores nos hacen clamar por una mayor cautela en las estrategias de inversión de los fideicomisos, en las que quizá se haga más hincapié en los bonos a corto plazo a fin de minimizar el riesgo de una inflación no prevista; con todo, el descenso sin precedentes en la tasa de bonos que orquestó el Sistema de la Reserva Federal a raíz de la crisis económica con la finalidad de aligerar la carga de la deuda y fomentar la inversión empresarial (al restarle todo atractivo a la inversión segura y pasiva por ser objeto de una remuneración escasa) arroja una sombra sobre esa respuesta a la volatilidad del mercado de valores.

XVI.10. Inversión social por los administradores fiduciarios

Algunas corporaciones hacen cosas que parecen inmorales en la opinión de segmentos influyentes de nuestra sociedad, y los críticos instan a los inversionistas a no comprar o conservar sus acciones. Si el inversionista es un individuo, su decisión de seguir o no ese llamado es puramente personal, una decisión que no plantea problemas interesantes para la política legal. Pero supongamos que el inversionista es el administrador de un fondo de pensiones o de una dotación universitaria y tiene las obligaciones de un administrador fiduciario ante los beneficiarios del fondo de inversión que administra. ¿Debiera permitírsele la práctica de la inversión "socialmente responsable", por ejemplo, expulsando de su cartera los valores de compañías que tienen una experiencia de discriminación sexual o racial, o, en el caso de un fondo de pensión para empleados municipales, invirtiendo en exceso (desde el punto de vista de los principios de inversión normales) en los valores de la ciudad?

[27] Stewart E. Sterk, "Rethinking Trust Law Reform: How Prudent Is Modern Prudent Investor Doctrine?", 95 *Cornell L. Rev.,* 851 (2010).

Primero, debe advertirse que el fondo incurrirá en gastos administrativos, semejantes a aquellos en los que incurren los seleccionadores de acciones, en identificar las empresas o las unidades gubernamentales cuyos valores se excluirán o se incluirán en exceso en la cartera; esto reducirá el rendimiento neto para los beneficiarios del fondo. Segundo, a menos que las empresas favorecidas y desfavorecidas traigan un empate al azar en el universo de inversiones posibles y además constituyan en conjunto sólo una pequeña fracción del total de las oportunidades de inversión, la inversión social conducirá a una diversificación insuficiente de la cartera del inversionista. Dado que, como sabemos, un inversionista no se ve compensado por mantener una cartera insuficientemente diversificada, el rendimiento esperado de la cartera no aumentará para compensar a los beneficiarios por soportar un riesgo mayor.

Sin embargo, adviértase que la teoría del financiamiento *no* predice que adoptar los principios de la inversión social conducirá a una disminución del rendimiento esperado para el inversionista, excepto por el monto de los costos administrativos adicionales en que se incurra. El rendimiento ajustado por el riesgo será menor si hay una diversificación insuficiente, pero el rendimiento (antes del gasto administrativo), omitiendo el riesgo, no será menor. No importa que las empresas favorecidas o desfavorecidas sean más o menos rentables que el promedio; los precios de sus acciones subirán o bajarán de tal modo que todas representen una inversión igualmente buena, juzgadas sobre una base de inversión por inversión antes que de toda la cartera, para el inversionista que siente aversión por el riesgo.

Dado que es improbable que los inversionistas sociales gasten mucho dinero en decidir cuáles valores excluirán (o incluirán excesivamente) en sus carteras, la consecuencia económica principal de la inversión social es la subdiversificación. Éste podría ser un problema particularmente grave en la inversión social de preferencia local. Cuando las ciudades de áreas deprimidas del país tienen dificultades para cumplir sus obligaciones de aportar dinero a los fondos de pensiones de los empleados municipales, los administradores fiduciarios de los fondos se ven presionados para que ayuden a la ciudad comprando sus valores. Si se rinden a la presión podrían causar una diversificación considerablemente insuficiente. Pero éste no es un análisis completo. El fondo deberá evaluar el trueque de ventajas y desventajas que tiene lugar entre la diversificación insuficiente y lo que podría ser una reducción de las aportaciones al fondo de pensiones si no hace un gesto de apoyo al municipio comprando sus valores en mayor cantidad que lo indicado como adecuado por la teoría financiera.

XVI.11. La regulación de los mercados de valores

La regulación del mercado de valores a través de la Comisión de Bolsa y Valores (SEC) data de la creación de esta última en el decenio de 1930 y tuvo como motivación la creencia en que el derrumbe de la bolsa de valores en 1929 había sido producto del fraude y otros abusos y, a su vez, había sido la causa de la Gran Depresión: *post hoc ergo propter hoc*.[28] Quienes se adhieren a la versión sólida (véase antes XVI.6) de la teoría de los mercados eficientes juzgarían mucho más probable que un declive precipitado en los precios de las acciones habría sido el resultado de la expectativa de un declive en la actividad económica en lugar de haber provocado el declive.[29] Pero, de hecho, la crisis económica que nos tocó vivir recientemente sugiere, como veremos, que las burbujas en los precios de los activos, las cuales no se basan en una anticipación de problemas económicos, pueden desencadenar una depresión o una recesión severa. Aun así, es de dudar que las provisiones clave de la legislación sobre los valores del decenio de 1930, tales como la exigencia de que las nuevas emisiones de valores se vendieran solamente por medio de un prospecto que debía presentarse por anticipado a la SEC para su revisión a fin de asegurarse de que contuviera toda la información que a ojos de esta dependencia es sustancial para los inversionistas, pudieran haber originado el desplome. Ciertamente esas disposiciones no precipitaron el desplome reciente, que vio cómo el Promedio Industrial Dow Jones pasó de estar por encima de 14 000 en 2007 a hundirse a menos de la mitad de ese puntaje en el invierno de 2009. Incidentalmente, tal hundimiento constituye un desafío adicional a la versión sólida de la teoría de los mercados eficientes: resulta poco plausible que a lo largo de ese periodo las ganancias futuras descontadas de las compañías en el Dow Jones cayeran a la mitad.

Escritos en una jerga legal y contable incomprensible, los prospectos no tienen ningún valor directo para el comprador de acciones no sofisticado (imaginemos la advertencia en un paquete de cigarrillos escrita con el vocabulario y en el formato de un prospecto de acciones). Tampoco es obvio que los requerimientos de revelación impuestos por la SEC aumenten en efecto el flujo de la información relevante transmitida al mercado. Al limitar los esfuerzos de venta a los prospectos y asumir una visión restrictiva de lo que podría incluirse adecuadamente en un prospecto (por ejemplo, la SEC se mostró durante largo tiempo hostil a las proyecciones de ganancias), la SEC ha limi-

[28] Para los antecedentes legales, véase Louis Loss y Joel Seligman, *Fundamentals of Securities Regulation* (5ª ed., 2003).

[29] Con todo, el desplome de la bolsa de valores puede haber contribuido a la gravedad de la Gran Depresión. Robert J. Gordon y James A. Wilcox, "Monetarist Interpretations of the Great Depression: An Evaluation and Critique", en *The Great Depression Revisited,* 49, 80 (Karl Brunner [comp.]., 1981).

tado la cantidad de información comunicada por los emisores. El peligro de que un esfuerzo en pos de acabar con la información tergiversada pueda también obstruir la divulgación de informes de índole genuina y provechosa se agudiza cuando esas noticias llegan a audiencias numerosas que conllevan costos heterogéneos en cuanto a la capacidad de absorber las nuevas; la información que puede desorientar a algunos miembros de la audiencia podrá ser valiosa para otros integrantes de ella. La revelación obligatoria puede también socavar el uso del secreto como un instrumento legítimo para la apropiación de los beneficios derivados del hecho de ser la primera compañía en hacer un descubrimiento valioso o en obtener una información comercialmente valiosa (véanse antes IV.7, XI.7).

Un estudio del economista George Stigler demostró que, antes de que se iniciara la regulación de las nuevas emisiones de valores o acciones en 1933, a los compradores de emisiones nuevas no les iba peor, en promedio, que a los compradores de emisiones nuevas en la actualidad.[30] Algunos economistas concuerdan con la conclusión de Stigler;[31] la evidencia empírica al respecto es débil y los supuestos teóricos en que se funda —la eficiencia de los mercados de valores— han sido cuestionados. El emisor tiene mucha mejor información acerca de su compañía que la que poseen incluso agentes externos de perfil sofisticado tales como los analistas de valores, y las ganancias potenciales para el emisor de subir el precio de la nueva acción son con frecuencia tan grandes como para dejar atrás las inquietudes relativas a la reputación que comúnmente inhiben un trato al instante y acaso inescrupuloso.[32] Un estudio cuidadoso encuentra que una ley sobre valores que demande la divulgación de información sustancial y provea de remedios legales de carácter privado para la renuencia a transmitirla, tal como la Ley de Valores de 1933, acarrea un mejoramiento neto en la operación de los mercados de valores, si bien formas más ambiciosas de regulación pública de los valores no lo hacen.[33] El primer enfoque, que posee una naturaleza de derecho común, es el predominante, como sería de esperar, en los regímenes jurídicos de origen inglés, y da evidencias en apoyo de la eficiencia superior del derecho común frente a los sistemas del derecho civil.

Pero ¿qué es exactamente lo que debe sacarse a la luz? ¿Debería hacerse pública la compensación otorgada a los directores ejecutivos de la empresa, por ejemplo? Probablemente no. Lo que sería necesario poner a la vista de

[30] George J. Stigler, "Public Regulation of the Securities Market", 37 *J. Bus.,* 117 (1964).

[31] Para un resumen de la literatura y pruebas adicionales, véase George J. Benston, *Corporate Financial Disclosure in the UK and the USA* (l976).

[32] Véanse Rafael La Porta, Florencio Lopez-de-Silanes y Andrei Shleifer, "What Works in Securities Laws?", 61 *J. Fin.,* 1 (2006); Michael Greenstone, Paul Oyer & Annette Vissing-Jorgensen, "Mandated Disclosure, Stock Returns, and the 1964 Securities Acts Amendments", 121 *Q. J. Econ.,* 399 (2006).

[33] La Porta *et al.*, antes nota 32.

todos es la información referida al valor de la compañía, y dicha compensación difícilmente se relaciona con ese valor: la compensación puede ser elevada porque la junta directiva no está protegiendo los intereses de los accionistas, o debido a que el director ejecutivo ocupa un lugar destacado.

Otro propósito de primer orden de la ley de valores (aunque éste es perverso) consiste en reducir la especulación, que en la negociación de valores como en otros renglones obedece a la saludable finalidad de facilitar el ajuste rápido de los precios a los valores actuales. (Recuérdese que el arbitraje es una forma de especulación.) El especulador está pronto a lanzarse en busca de valores subvaluados para comprarlos y de valores sobrevaluados para venderlos en corto. La información que extrae de la sima en que yacía se difunde velozmente a través del mercado (la velocidad con la que llega a todos los rincones la información sobre valores negociados en intercambios organizados es la razón principal de que sea difícil superar en desempeño sistemáticamente al mercado de valores), lo que así les permite a otros negociantes ajustarse con prontitud a los cambios en las condiciones que han descubierto.[34]

La ley discrimina contra quienes especulan en torno a una recesión en el mercado, prohibiendo que las acciones se vendan en corto a un precio menor que el precio más reciente al que se ha negociado la acción. La actitud legal se asemeja a la antigua práctica de castigar a un portador de malas noticias. Quien vende una acción en corto —conviene en entregar al precio corriente del mercado, esperando que el precio de la acción bajará de modo que pueda comprarla cuando deba hacerse la entrega por menos de lo que la ha vendido— perderá dinero si no ha pronosticado correctamente una disminución del precio de la acción. Pero no puede producir esa baja a menos que propale falsos rumores, lo cual es ilegal. La actitud hacia la venta en corto es especialmente incongruente en vista de la preocupación por evitar pánico. En la medida en que se desaliente esa venta, es probable que las declinaciones del mercado sean más marcadas, no menos.

Otro esfuerzo para disminuir la especulación es limitar la compra de acciones en el margen, es decir, del monto del apalancamiento que podrá emplearse en la compra de acciones. Es poco probable que este esfuerzo tenga éxito. El apalancamiento es solamente un procedimiento para aumentar el riesgo y el rendimiento esperado de una compra. Una alternativa a la que no alcanza la limitación del margen consiste en mantener una cartera de valores de alto riesgo. Una demanda de tales valores llevará a las corporaciones a aumentar el monto del apalancamiento en sus estructuras de capital, porque ello aumentará el beta de sus acciones. Como veremos en breve, la toma de

[34] Los especuladores ni siquiera ganan (en promedio) rendimientos elevados; en consecuencia, la sociedad compra estos servicios valiosos a bajo precio. Lester G. Telser, "Why There Are Organized Futures Markets", 24 *J. Law & Econ.,* 1, 9-10 (1983).

riesgos excesivos puede generar un severo riesgo a escala macroeconómica, pero sencillamente regular el margen equivale a tapar el sol con un dedo.

La ley federal de valores es también un vehículo para dotar de jurisdicción a los tribunales federales sobre los fraudes claros con valores. La Regla 10b-5 de la SEC, que se ha interpretado en el sentido de que crea un derecho a demandar por daños, prohíbe el fraude en la compra o la venta de valores y así permite que los inversionistas litiguen reclamaciones de fraude en el tribunal federal.

Por esta razón, resulta vital saber qué es un valor y qué no lo es. Supongamos que una compañía aseguradora vende una anualidad que garantiza al tenedor un rendimiento mínimo sobre su inversión, pero el rendimiento mayor dependerá del desempeño de la cartera de valores en la que la compañía aseguradora invierte las aportaciones que él haga. ¿Es la anualidad un valor? Lo es (dicen los tribunales), porque el rendimiento del inversionista es en una medida considerable una función del desempeño de lo que irrefutablemente son valores.

Pero supongamos ahora que el propietario de un negocio incorporado lo vende a otra corporación. El comprador compra todas las acciones en camino hacia la fusión de la empresa adquirida. Si piensa que el vendedor hizo declaraciones falsas en la venta, ¿podrá demandar en los términos de la Regla 10b-5 como la víctima de un fraude con valores? Podrá hacerlo. Este resultado tiene escaso sentido económico. Aunque el comprador está comprando indiscutiblemente valores, no necesita las protecciones que las leyes de valores otorgan a los inversionistas pasivos, porque no tienen el interés o (a menudo) la experiencia que los motivaría o los capacitaría para protegerse a un costo razonable. Nadie compra todo un negocio sin una investigación cuidadosa. Tal comprador no necesita protecciones legales fuera de las que le otorga el derecho común de los fraudes.

Veamos ahora los problemas de los daños en los casos de fraude con valores. Según los principios del fraude convencional, una declaración falsa, para que genere una acción, debe haber sido la base de la decisión de la persona presuntamente defraudada; de otro modo, el fraude es inocuo. Supongamos entonces que una declaración falsa en un prospecto de una emisión nueva de acciones lleva a los corredores de la bolsa que lean el prospecto a comprar grandes cantidades de la acción y a recomendar a sus clientes que hagan lo mismo. En consecuencia, el precio aumenta. Supongamos que alguien que no conozca el prospecto —que en efecto no tenga idea de la causa por la que el precio de la acción ha aumentado— la compra a su precio mayor. Más tarde se descubre el fraude y el precio baja. ¿Debiera permitirse que este inversionista demandara al emisor? La Corte Suprema ha contestado afirmativamente,[35]

[35] Basic *vs.* Levinson, 485 U.S., 224 (1988); y para una discusión reciente de carácter amplio, Bowe *vs.* Po1yMedica Corp., 432 F.3d, 1 (lst Cir., 2005).

y ésta es la respuesta económicamente correcta. El fraude está implícito en el precio del mercado, y la persona que compra sin conocimiento del prospecto está actuando con base en la información falsa en la misma medida que quienes compran con conocimiento.

Pero consideremos ahora cuál debiera ser el *monto* de los daños. A primera vista, pareciera ser obvio: deben ser las pérdidas de quienes compraron a un precio inflado por el prospecto falso u otra declaración tergiversada, tras la eliminación (mediante el modelo de precios de los activos de capital bosquejado antes, en XVI.1) de otras causas posibles de la baja del precio de la acción. ¿Pero qué diremos de las personas inocentes que se beneficiaron con el fraude? Supongamos que una persona engañada por el prospecto compra la acción cuando su precio está aumentando, pero se deshace de ella con una ganancia antes de que explote la burbuja. A menos de que se le obligue a devolver su ganancia, la corporación que emitió el prospecto engañoso será obligada a pagar daños en exceso del daño causado a los compradores defraudados. En virtud de que no hay ninguna base legal o práctica para obligar a la restitución de manos de los accionistas que fueron injusta pero inocentemente enriquecidos por el fraude, existe un peligro de disuasión excesiva (compárense antes VI.7, X.14, donde discutimos problemas similares en la ley de los cuasidelitos y contra los monopolios).

Podemos hacer una afirmación más fuerte, ¿o no? A menudo, las ganancias derivadas de un fraude con acciones serán nulas, por lo menos tal como se miden de ordinario los daños en los casos de fraude. Supongamos que los administradores de una corporación demoran injustificadamente la revelación de malas noticias acerca de las perspectivas de la corporación, con la vana esperanza de que algún milagro restaure tales prospectos. En esta forma, los administradores posponen por dos semanas un hundimiento del precio de las acciones de la corporación. Quienes compren la acción durante este periodo serán perjudicados, pero los vendedores se beneficiarán porque si hubieran conservado la acción durante más tiempo habrían sufrido la pérdida que en su lugar sufrieron los compradores. Si los propios administradores se hubiesen beneficiado de la venta de su acción antes de que las malas nuevas llegaran al mercado, no habría ningún problema para obligarlos a devolver sus ganancias (es decir, las pérdidas que evitaron). Pero supongamos que no lo hicieron. Entonces, ¿cuál sería la base para el otorgamiento de daños basado en las pérdidas de algunos de los accionistas? Una base pragmática es la siguiente: si se la obliga a pagar las pérdidas, la corporación tendrá un incentivo para vigilar a sus administradores con mayor cuidado en el futuro.

Todo esto supone que hay *algún* costo social del fraude, aunque no un costo igual a las pérdidas de los accionistas que se vieron perjudicados por el ocultamiento de los administradores. Son posibles dos clases de costo social, aunque sean difíciles de cuantificar. Primero, es posible que los administradores hayan gastado recursos reales con el ocultamiento de las malas nuevas.

Segundo, y relacionado con lo anterior, algunos inversionistas tendrán ahora un incentivo para gastar más recursos tratando de descubrir la verdad acerca de las empresas, porque ahora tienen que superar los esfuerzos de desinformación de los administradores corporativos. Sin embargo, estos costos podrían ser mucho menores que las ganancias del juicio (para los inversionistas perjudicados por el fraude). Aun así, no se sigue de aquí que habrá demasiados juicios desde un punto de vista social, si la amenaza del juicio impide los fraudes en gran medida.

Supongamos que el prospecto no revela que uno de los principales funcionarios del oferente tiene antecedentes penales, y esta omisión es importante en el sentido de que el demandante no habría comprado la acción si hubiese sabido de aquéllos. Es una acción petrolera, y más tarde se derrumba el precio del petróleo y en consecuencia la acción pierde valor. En un juicio por declaraciones falsas, ¿debiera la pérdida de valor ser un componente de los daños?[36] La respuesta es no. El inversionista quería una acción de petróleo, así que si no hubiera comprado ésta habría comprado otra, según las probabilidades, y habría sufrido la misma pérdida.

XVI.12. UNA REVISIÓN DE LA COMPENSACIÓN: EL CASO DE LAS COMISIONES DE LOS FONDOS MUTUOS

La sección 36 (b) de la Ley de Compañías de Inversión[37] hace que el asesor de fondos mutuos se convierta en el agente fiduciario de los fondos mutuos sobre los que brinda consejo. El Segundo Circuito, en el caso Gartenberg *vs.* Merill Lynch Asset Management, Inc.,[38] estableció que al decidir si un asesor de fondos mutuos ha faltado a su obligación fiduciaria puede sopesar la posibilidad de que el cobro sea "tan desproporcionadamente alto que no guarda ninguna correspondencia razonable con los servicios prestados y no pudiera haber sido el producto de una negociación realizada a suficiente distancia". Muchos años después el Séptimo Circuito desaprobó el Gartenberg, aseverando que "un administrador fiduciario debe revelar las cosas a plenitud y no hacer ningún truco, pero no está sujeto a que se le ponga un tope a su compensación".[39]

[36] Véase Bastian *vs.* Petren Resources Corp., 892 F.2d, 680 (7th Cir., 1990). La regulación de los mercados de futuros (el intercambio de futuros se discutió brevemente en el capítulo IV) plantea también interesantes problemas económicos. Los artículos en 4 *J. Futures Mkts.,* 3 (otoño de 1984) proveen un buen punto de partida para el lector que desee explorarlos. Véase también Frank H. Easterbrook, "Monopoly, Manipulation, and the Regulation of Futures Markets", 59 *J. Bus.,* S103 (1986).

[37] 15 U.S.C. 80a-1 *et seq*.

[38] 694 F.2d, 923, 928 (2d Cir., 1982).

[39] Jones *vs.* Harris Associates L.P., 527 F.3d, 627, 632 (7th Cir., 2008), en libertad bajo fianza, 130 S. Ct., 1418 (2010). La Corte Suprema expresó su acuerdo general con el enfoque que el Segundo Circuito había adoptado en Gartenberg.

El asesor en el caso del Séptimo Circuito había fundado y administrado algunos de los fondos mutuos sobre los que prestaba asesoría; otros de sus asesorados eran independientes. Cobraba el doble por el servicio de asesoría a sus fondos "cautivos". El Séptimo Circuito afirmó que "el hecho de que los fondos mutuos sean 'cautivos' de sus asesores de inversión no restringe esta competencia. Un asesor no puede hacer dinero a base de su fondo cautivo si sus honorarios elevados ahuyentan a los inversionistas".[40] Pero, en apariencia, los honorarios altos no habían hecho que salieran corriendo los inversionistas, pese a que hubiera salido a la luz que el cobro era elevado. Por supuesto, una posibilidad es que los fondos habían funcionado de maravilla, con todo y los honorarios elevados. Aun así, se supone que la compensación recibida por un administrador fiduciario no está basada en el costo sino en la ganancia; se supone que le corresponde ser el *alter ego* de sus principales. Es difícil comprender por qué al asesor le podía costar el doble administrar los fondos que había creado de lo que le costaba hacer lo mismo con los fondos sobre los que no tenía ningún control. ¿Por qué no se habían percatado de ello los inversionistas y no se habían quejado? ¡Una razón podría ser que se tomaban muy en serio el estatus fiduciario del asesor!

XVI.13. ¿Por qué regular a los bancos?

Los bancos son intermediarios financieros. En otras palabras, toman prestado de los dueños de capital y prestan el dinero obtenido a personas o compañías que requieren capital. La intermediación financiera (las actividades bancarias en un sentido amplio que no está limitado a la actividad de los bancos comerciales) es de forma inherente riesgosa. El banco necesita cobrar a las personas a quienes les presta una tasa de interés más alta que aquella a la que toma en préstamo, a fin de que pueda pagar salarios y cubrir otros gastos y proveer de un rendimiento a sus accionistas. Lleva a cabo esto tomando un préstamo en corto y prestando en largo. (Es esto lo que recibe el nombre de "arbitraje de los vencimientos".) Una tasa de interés es más baja (siendo todas las demás circunstancias iguales) entre más breve sea el plazo del préstamo. El prestamista retiene una liquidez considerable (efectivo o su equivalente) cuando presta por solamente un periodo breve), y la liquidez es valuada como una cobertura contra la incertidumbre; y el riesgo de incumplimiento es más bajo entre más breve sea el plazo, dado que es menos probable la ocurrencia de situaciones imprevistas que pudieran generar un impago. En la misma lógica, entre más largo sea el plazo, más alta será la tasa de interés, con el objeto de compensar al prestamista por la pérdida de liquidez y el riesgo más alto de incumplimiento.

[40] 527 F.3d, en la 632.

Debido a que las tasas de préstamo a corto plazo son más bajas no sólo que las tasas de préstamo a largo plazo sino también que los rendimientos de capital social (en razón de la mayor incertidumbre que rodea a los valores de capital), los bancos prefieren tener su capital dominado por capital a corto plazo que por capital social; de ahí que se empeñen en fijar una razón muy alta de deuda a capital social. No obstante, este modelo de negocios coloca al banco en el riesgo de tener que satisfacer demandas de que la mayor parte de su capital dé rendimientos (a los prestamistas a corto plazo que suministran la mayoría del capital en manos de la institución) mientras su capital permanece ligado a préstamos a largo plazo. Supongamos que uno de los prestamistas del banco (sus proveedores de capital) descubre que algunos de los préstamos son precarios y, por consiguiente, ese prestamista exige la devolución de su capital (su depósito, de acuerdo con la nomenclatura convencional de la banca comercial). Otros depositantes, al enterarse de que el primer depositante ha ordenado el retiro, pueden, ya sea porque están de acuerdo con su evaluación o porque temen que otros depositantes lo estén, decidirse a retirar igualmente su capital. Presa de la desesperación por la necesidad de satisfacer estas demandas, el banco puede tratar de vender algunos de sus préstamos, pero su desesperación puede desembocar en que tenga que vender a un precio demasiado bajo para generar los fondos que necesita a fin de hacer frente a todas las exigencias de retiro; y a continuación cae en quiebra. Supongamos para efectos de simplicidad que el banco tiene un capital en depósito por un total de 100 dólares, pero solamente 5 dólares en efectivo, mientras que los 95 restantes están ligados a préstamos. Si la mitad de los depositantes pidieran de vuelta su capital, el banco tendría que encarar el asunto recurriendo a un efectivo adicional de 45 dólares, lo que requeriría que liquidara casi la mitad de su cartera de préstamos a la mayor brevedad.

El seguro federal de depósitos se instituyó en la década de 1930 con el propósito de prevenir las estampidas bancarias. La idea era que un depositante asegurado no debía preocuparse por el colapso de su banco si su depósito no excedía el límite asegurable. Percibirá el paralelo entre el seguro de depósitos y la Ley de Quiebras. Dado que un depositante tiene un reclamo pagadero en caso de demanda, podría, en ausencia de un seguro de depósitos, retirar su depósito al menor signo de que su banco se halla en problemas financieros. Otros depositantes seguirían sus pasos en un santiamén, o de lo contrario les tocaría pagar los platos rotos. Los esfuerzos que dirigieran todos los depositantes a efectuar un retiro de inmediato o en un plazo corto traerían por consecuencia la quiebra del banco. Una estampida bancaria es un ejemplo clásico de una externalidad dañina; al hacer un intento por retirar su depósito, cada depositante infligirá un daño a los demás, pero al sopesar su decisión no tomará en cuenta la lesión que está causando.

La decisión del gobierno de asegurar los depósitos bancarios determinó la inevitabilidad de que se efectuara alguna clase de supervisión reguladora

debido a que cada asegurador, público o privado, posee un interés en controlar (si eso es viable) los riesgos contra los que proporciona un seguro, en especial cuando éstos se hallan bajo el control del asegurado. Una forma de conceptualizar a las agencias bancarias federales es como sustitutas de los depositantes cuya función es insistir en los tipos de protección al acreedor que los depositantes exigirían si no estuvieran asegurados. Y, de hecho, muchas regulaciones bancarias se asemejan a los tipos de protección en los que los acreedores privados insisten a menudo. Ejemplos de esto son los requisitos de capitalización mínima de los bancos asegurados y los límites fijados al carácter riesgoso de las carteras de préstamos que ofrecen los bancos asegurados. Los requisitos de capitalización mínima reducen el riesgo de un acreedor asegurando que exista un colchón de capital accionario de un tamaño prescrito, y la limitación fijada a los riesgos que un prestatario puede imponer a sus activos impide que éste reduzca de forma unilateral la tasa de interés negociada (ajustada al riesgo).

Por añadidura, se prohibió a los bancos que pagaran interés sobre los depósitos en demanda. Al otorgar a los bancos una fuente de capital que estaba exenta de intereses (los bancos tenían poder monopsónico sobre los depósitos cuando no existían sustitutos cercanos al depósito en demanda asegurado federalmente), esta restricción facultó a los bancos para hacer préstamos a corto plazo que eran seguros y aun así también lucrativos —aun cuando la tasa de interés sobre los préstamos a corto plazo es baja— porque el costo para el banco del capital que éste usó a fin de realizar esos préstamos era de cero.

La regulación bancaria no se detuvo allí. Una manera de reducir la probabilidad de que un banco entre en quiebra y así traiga como efecto inmediato la obligación del gobierno de rembolsar a los depositantes asegurados es restringir la competencia entre los bancos (algo que los mismos bancos anhelan). La prohibición levantada contra el pago de intereses sobre las demandas en depósito era una limitación impuesta a la competencia de precios entre los bancos. También se limitó el acceso a actividades bancarias, tanto directamente como a través de interdicciones sobre las sucursales bancarias. La Ley Glass-Steagall (hoy en día revocada) prohibía a los bancos involucrarse en la suscripción de valores. Si bien se consideraba que esa actividad era riesgosa, el verdadero propósito de la Ley era probablemente el de proteger a los suscriptores de valores de la competencia de los bancos. Asimismo, prevalece una fuerte sospecha de que la razón genuina para fijar límites a la creación de sucursales bancarias y de bancos interestatales no era proteger a los bancos contra los riesgos de la competencia, que podrían ser encarados a final de cuentas por las agencias federales de seguros de depósitos, sino proteger a los bancos de la competencia, y punto.

A continuación emergieron los fondos del mercado monetario, los cuales ofrecían un sustituto cercano de cuentas de cheque de los bancos pero pagaban un interés, con lo que atrajeron depósitos de los bancos. Y tiempo más

tarde los corredores de la bolsa (empresas involucradas en la correduría y la negociación de bonos y acciones), tales como Goldman Sachs y Morgan Stanley, empezaron a competir en mercados de préstamos con los bancos comerciales utilizando capitales a corto plazo que, sin embargo, no estaban asegurados federalmente. A lo largo de un periodo de dos décadas que se abrió a comienzos de 1980, viéndose forzados a enfrentar la competencia de los fondos de mercado monetario y de los corredores de la bolsa ("bancos no bancarios"), y teniendo que navegar bajo la marea de la desregulación (abordada en el capítulo XIII), los bancos comerciales quedaron liberados de muchas de las restricciones de la era de la década de 1930 que pesaban sobre su libertad para competir.

Entre las medidas desregulatorias estaba incluida una ley que permitía a las asociaciones de ahorros y préstamos (las cuales se especializan en préstamos hipotecarios y ahora son conocidas como "cajas de ahorros") ofrecer por vez primera cuentas de cheques haciéndoles la competencia a los bancos comerciales. A finales de la década de 1980 el desastre se había abatido contra las asociaciones de ahorros y préstamos, que fueron cayendo en quiebra una tras otra. El principal culpable de ello, según parece, fue la ausencia de evaluación de la experiencia por parte de los seguros federales de depósito. La evaluación de la experiencia, como sabemos (véase antes VI.14), significa establecer una variación entre las primas de seguro dependiendo del riesgo (de impago, accidente o cualquier otra calamidad contra la que se esté asegurado) que afronta la entidad particular asegurada, en vez de cobrar una prima uniforme a todos los asegurados. Una vez que fueron de nuevo autorizadas a competir por los depósitos pagando un interés, las asociaciones de ahorros y préstamos tenían un incentivo para buscar inversiones con alto rendimiento, dado que eso les permitiría pagar tasas de interés más elevadas y de esta manera crecer. Sabemos que tener un rendimiento esperado que es alto significa un riesgo elevado, pero, debido a que el seguro federal de depósito no se guía por las evaluaciones de la experiencia, dichas asociaciones podían desentenderse de esos riesgos y poner la carga en las espaldas de los contribuyentes. Por supuesto, los accionistas tuvieron que sobrellevar un riesgo añadido, pero pudieron evadirlo a base de diversificaciones gracias al diseño apropiado de sus carteras; con todo, los administradores no podían hacer idéntica cosa con ningún capital humano en su haber que fuera específico de la industria de ahorros y préstamos. Los reguladores no se toparon con el problema sino hasta demasiado tarde. Con la evaluación de la experiencia realizada por los seguros de depósito se reduciría la probabilidad del fracaso generalizado de las instituciones financieras sin incurrir en los costos sociales de la reregulación.

XVI.14. EL DESPLOME DE 2008

El revocación de la Ley Glass-Steagall en 1999 completó en gran medida la desregulación de la industria bancaria. No por ello puede afirmarse que esta industria haya quedado absolutamente desregularizada. En efecto, las agencias bancarias, sobre todo la Reserva Federal y la Comisión de Bolsa y Valores, retuvieron poderes extensos para regular las prácticas y las estructuras de capital de los bancos comerciales y los corredores de bolsa, respectivamente. Pero durante la administración de George W. Bush estos poderes se ejercieron en forma debilitada. La mayoría de los economistas y legisladores, así como los integrantes de la Reserva Federal y los funcionarios económicos de la administración, abrigaban la creencia de que, aunque la actividad bancaria se había vuelto más riesgosa de lo que era antes de la desregulación, podía depositarse la confianza en los banqueros en cuanto que para ellos era un asunto de interés propio evitar el tomar riesgos excesivos, puesto que entendían a fondo la naturaleza riesgosa que de por sí implicaba su negocio.

Y, si las agencias no prevenían las quiebras de los bancos, ¿qué importaba? La industria aérea había sido objeto de la desregulación sin que ello atrajera malas consecuencias, aun cuando, a semejanza de las actividades bancarias, constituye un negocio inherentemente riesgoso y, a resultas de esta peligrosidad, a menudo se halla al borde de la quiebra e incluso en ocasiones se lanza al precipicio. Las líneas aéreas entrañan un riesgo porque conllevan costos fijos muy pesados, una fluctuación en la demanda y, asimismo, una fluctuación en los costos de insumo (fundamentalmente, combustible para la aviación). Una industria cuyos costos fijos son onerosos tiende, como sabemos, a ser inestable, y en particular si sus otros costos y sus ingresos fluctúan sobremanera. Después de los ataques terroristas del 11 de septiembre de 2001, la industria de las aerolíneas estadunidenses se encontraba esencialmente en bancarrota (tuvo lugar un rescate modesto y ahora olvidado que el gobierno ideó con la intención de negarle a Osama bin Laden la satisfacción de haber hecho quebrar a una industria de los Estados Unidos), puesto que sus ingresos se hundieron en tanto que sus costos fijos permanecieron… fijos. Al año siguiente, United Airlines en efecto se declaró en quiebra. No obstante, continuó estando en la esfera comercial, y el impacto de las penalidades financieras de las aerolíneas sobre el resto de la economía nacional y mundial fue ligero.

Lo que se pasó por alto fue que una crisis de solvencia en la industria bancaria, a diferencia de una crisis de solvencia en la industria de las aerolíneas, puede representar una fuente de enormes costos externos, lección que ahora nos hemos aprendido. La mayoría de las personas que quedaron desempleadas a consecuencia de la desaceleración económica que sobrevino tras el colapso financiero de 2008 no tenía nada que ver con la industria bancaria. Aun cuando los accionistas y los ejecutivos de una serie de bancos de considera-

ble magnitud perdieron dinero, sus pérdidas fueron mucho más pequeñas que las sufridas por la economía estadunidense en su conjunto.

La razón de que una crisis bancaria, a diferencia de una crisis en las aerolíneas, pueda imponer costos externos sustanciales es que los bancos (y recuérdese que esta palabra es una forma abreviada para referirse a los intermediarios financieros, y por ello no se limita a los bancos comerciales) son una fuente crucial de crédito para la economía. Los consumidores y los negocios por igual piden prestado dinero para cubrir el intervalo que se abre entre el momento en que incurren en gastos y aquel en que reciben un ingreso para pagarlos; si no pueden obtener un préstamo, ya sea porque los bancos son insolventes o temen volverse insolventes y, por consiguiente, no pueden hacer o temen hacer préstamos, la actividad económica se derrumba estrepitosamente. El derrumbe puede inducir a una peligrosa espiral descendente en la economía. Los consumidores escatiman el gasto, lo que conduce a los productores a disminuir la producción y la venta y, de forma concomitante, a despedir trabajadores, quienes a su vez tienen menos gastos debido a que sus ingresos y sus perspectivas de ingresos han menguado, lo que desemboca en una caída ulterior en la producción y un incremento en el desempleo. En un escenario en el que los precios decrecen a medida que los productores compiten entre sí desesperadamente para conservar a sus clientes, cada vez provistos de menos recursos, la gente comienza a atesorar efectivo con la expectativa de que los precios bajen aún más; y este atesoramiento, que indica una espiral deflacionaria, provoca una caída adicional en el gasto del consumidor y, por ende, en la producción.

La incertidumbre constituyó un factor de peso en la desaceleración económica. Una vez que el crédito se congeló en septiembre de 2008, nadie podía estar seguro de qué tan aciagas podrían ser las consecuencias de la medida. Una respuesta racional del ser humano a la incertidumbre (véase antes I.1) consiste en quedarse paralizado, con la esperanza de tener la capacidad para disipar la incertidumbre o atenuarla antes de que le toque actuar, de manera que pueda actuar de modo más inteligente. (Es algo parecido a comprar una opción.) Pero cuando los consumidores se quedan paralizados el consumo se abate; y, conforme los negocios tienen menos ventas y recortan los proyectos de inversión, el empleo se viene abajo y con él el consumo.

¿Qué fue lo que desencadenó la crisis de solvencia de la industria bancaria? Los factores clave parecen haber sido la endeble política monetaria que se adoptó a principios de la década de 2000 y la falta de atención regulatoria que se siguió de largo hasta el desplome de septiembre de 2008. Animada por el deseo de estimular la economía luego del colapso subsiguiente a la burbuja de las empresas .com, la Reserva Federal, bajo las órdenes de Alan Greenspan (su director por largo tiempo), dispuso un descenso considerable de las tasas de interés a corto plazo y las mantuvo en ese nivel por varios años; durante una gran parte de este periodo fueron de hecho negativas, dado que la tasa

inflacionaria era más elevada que ellas.[41] Aunque las tasas de interés hipotecario son a largo plazo, una reducción en las tasas a corto plazo tiene probabilidades de reducir aquéllas y de hecho las redujo, debido a que los préstamos a corto y a largo plazo son recíprocamente sustituibles. Por ejemplo, un préstamo a corto plazo sustituible del que se espera que sea extendido cuando expire es (dependiendo del número de extensiones) un equivalente de un préstamo a largo plazo. De hecho, la tasa de interés sobre un préstamo a largo plazo se puede calcular como la suma descontada de las tasas de interés sobre una serie de préstamos a corto plazo a lo largo del mismo periodo. La tasa de interés hipotecario, si bien por lo común es a largo plazo, sufrió un descenso al efectuarse una reducción en las tasas de interés a corto plazo.

Dado que las casas se compran mediante una deuda considerable (tradicionalmente una hipoteca de 80%, pero a menudo en años recientes con un porcentaje de 100%), una reducción en las tasas de interés hipotecario reduce el costo de la vivienda y así incrementa la demanda. Los precios de la vivienda, que ya se estaban elevando raudamente, siguieron creciendo a principios del decenio de 2000 con el auxilio de tasas hipotecarias muy bajas. No es que los precios se eleven siempre que aumenta la demanda. Ello depende de la pendiente que se observe en la curva de la oferta. Si es horizontal, una demanda acrecentada se traduce en una producción acrecentada al mismo precio. Pero el valor de la vivienda es extremadamente durable, y, a causa de demoras en la creación de nuevo parque inmobiliario, la oferta de nuevas viviendas a corto plazo es inelástica. Como resultado, un aumento intensificado en la demanda de viviendas desemboca en precios más altos y no solamente en más construcción de casas y conjuntos de departamentos.

Llegado cierto punto, conforme los precios de las viviendas siguen aumentando, se instaura el comercio movido por el ímpetu: las personas vieron cómo los precios seguían subiendo y decidieron sumarse al jolgorio incorporándose al mercado, o si ya eran propietarios de una casa decidieron venderla y comprarse una mejor o conseguir un préstamo inmobiliario de capital social y utilizar los beneficios para comprar otros bienes o servicios. Si los precios de algún tipo de activos están incrementándose, entonces, aun cuando nadie sabe por cuánto tiempo persistirá la tendencia al alza de los precios, el solo hecho de que muchos inversionistas estén comprando (lo que en el contexto de una burbuja constituye *el motivo* por el que los precios están aumentando) es un indicio de que algunos de ellos, al menos, piensan que el activo es una ganga (la demanda creciente de cara a los precios crecientes sugiere que entre la gente predomina la noción de que el activo está subvaluado en

[41] La Reserva Federal manipula tasas de interés a corto plazo, en efecto, al comprar y vender valores del Tesoro a corto plazo. Si compra un valor incrementa el suministro de dinero en la economía, lo que estimula la actividad económica; si lo vende y retira el efectivo que recibe, reduce el suministro de dinero y de esta forma lentifica la actividad económica. Entre más dinero provee, menor es la tasa de interés, ya que es mayor el suministro de dinero para préstamo.

su precio) y lo seguirán comprando, en todo caso por cierto tiempo, si bien otros pueden estarlo comprando sólo por creer que *los demás* inversionistas lo seguirán comprando. Así pues, los precios se elevan porque los precios se elevan y no debido a que haya habido cambios en los principios básicos de la demanda y la oferta. Éste es el fenómeno de la burbuja de precios de los activos, una consecuencia no poco común del comercio movido por el ímpetu.

La industria bancaria dio acogida a la demanda creciente al facilitar las condiciones hipotecarias; las hipotecas de tasas ajustables le permitieron a la gente comprar casas sin un pago inicial o de anticipo y requerían sólo tasas "tentadoras" por lo bajas para los primeros dos o tres años de la hipoteca, de modo que si los precios continuaban subiendo el prestatario tendría un capital accionario en su casa al momento de expiración de las tasas tentadoras y podría refinanciar con una hipoteca convencional de 30 años a una tasa de interés baja. Las hipotecas, a menudo por miles a la vez, se empacaban en grandes bonos ("valores respaldados por hipotecas"). El ingreso generado por las hipotecas constituía el ingreso del bono. En vez de vender el bono entero (que podría valuarse en mil millones de dólares o más), su creador lo dividía luego en "pedazos" (rebanadas) que tenían diferentes prioridades y conllevaban distintas tasas de interés. Por ejemplo, el pedazo más antiguo podría darle a quien lo comprara el derecho al primer 80% de cualquier ingreso generado por las hipotecas. El comprador de este pedazo estaría en buena situación siempre que las hipotecas efectivamente dieran un rendimiento de 80% del principal y el interés debidos por los prestatarios (los deudores hipotecarios) y, por ello, este comprador tendría prometida sólo una tasa de interés modesta por parte del emisor del valor respaldado por la hipoteca. Los compradores de pedazos recién formados afrontarían un riesgo mucho mayor y, por lo tanto, tendrían prometidas unas tasas de interés más altas en compensación.

La valorización de las hipotecas significó un incremento en la demanda de inversión en las hipotecas. Y es que resulta mucho más conveniente poseer un bono que una hipoteca. La hipoteca requiere servicio y potencialmente una demanda de cierre; el bono no. El servicio y otra forma de administrar las hipotecas que subyacían a los valores con respaldo hipotecario quedaron a cargo del emisor de los bonos que representaban intereses en esas hipotecas; los compradores no tenían responsabilidades administrativas y podían vender los bonos sin mayores esfuerzos. Así, los valores respaldados por hipotecas succionaron el capital derivándolo hacia el financiamiento de las hipotecas y, de esta manera, contribuyeron a que la burbuja de la vivienda se mantuviera boyante. Aun cuando muchas de las hipotecas eran de alto riesgo (esto es, el prestatario no contaba con buen historial crediticio), los valores con respaldo hipotecario estaban diversificados geográficamente. Y el procedimiento del corte de pedazos les permitió a los inversionistas adversos al riesgo alcanzar una protección adicional contra el impago comprando justamente el pedazo más antiguo. Con todo, el hundimiento en

los precios de la vivienda ocurrido en 2007 y 2008, después de que estallara la burbuja, fue tan inesperadamente abismal (alrededor de 30% en toda la nación) que incluso los compradores de pedazos antiguos perdieron su dinero en virtud de que tantos deudores hipotecarios incumplieron.

Cuando los precios en la burbuja de precios de los activos llegan al punto donde casi todo aquel que desea y puede adquirir el activo lo ha comprado, el precio se estabiliza en una meseta (fue lo que ocurrió en el mercado de la vivienda en mayo de 2006) y a continuación es bastante probable que caiga mientras las personas que compraron recientemente el activo con la esperanza hoy hecha añicos de que su precio se iba a elevar comienzan a pedir rescate. Y ahora el ímpetu echa reversa. Al ver a la gente pidiendo rescate, otros dueños del activo se preocupan de que el valor de mercado del activo haya de caer y, en consecuencia, comienzan a efectuar la venta, por lo que empujan al precio hacia abajo. La prevalencia de las hipotecas de tasas ajustables aceleró esa caída. Por lo general, esas hipotecas implican tasas de interés muy bajas para los primeros dos años, lapso después del cual la tasa de interés es reconfigurada a un nivel mucho más alto. El deudor hipotecario (prestatario) puede haber tenido la certeza de que sería capaz de refinanciar la hipoteca a una tasa más baja (porque esperaba que su casa aumentara de valor, con lo que se incrementaría el capital accionario implícito en ella y, por ende, el valor que le podría ofrecer al prestamista). Cuando los valores de las viviendas se fueron a pique el refinanciamiento se volvió poco factible, y a menudo el deudor hipotecario se encontró con que la vivienda tenía un capital accionario de signo negativo, lo que podría llevarlo a incumplir.

Uno de los bancos que habían invertido en los valores con respaldo hipotecario era Lehman Brothers, un corredor de bolsa. Estaba sumamente apalancado, y la mayor parte de su capital era a largo plazo, y, en virtud de que era un "banco no bancario", su capital no estaba asegurado federalmente. Cuando el valor de mercado de los valores con respaldo hipotecario cayó estrepitosamente, la gente acudió corriendo en tropel a Lehman; y, cuando el gobierno se rehusó a rescatarlo, de manera que se declaró en quiebra, las estampidas bancarias se dirigieron igualmente a otros corredores de bolsa, y en este punto la industria crediticia del mundo entero comenzó a paralizarse. Los precios de los activos financieros se desplomaron mientras los bancos liquidaban sus carteras de inversión desordenadamente para satisfacer las demandas de retiros de capital. El gobierno intervino con préstamos e inversiones masivos en la industria financiera, pero era ya demasiado tarde para evitar una crisis económica global hecha y derecha.

Los banqueros han recibido críticas por no haber anticipado e impedido el colapso financiero. La crítica es errada. No es racional para una empresa conducir su negocio de forma tal que se reduzca a cero el riesgo de bancarrota. El peligro de quiebra es un costo esperado de hacer negocios, como cualquier otro costo, y una empresa racional elevará los riesgos hasta el punto en que

el costo esperado sea igual justamente a la ganancia obtenida por asumir los riesgos. La mayoría de los economistas, funcionarios de gobierno e inversionistas privados (incluido Warren Buffett), creían que el riesgo de un colapso de la industria bancaria era insignificante. Esto hizo que fuera prácticamente imposible para cualquier banco de gran magnitud adoptar un modelo de negocios que redujera en grado significativo el riesgo de quiebra. Recuérdese que el capital bancario es principalmente de corto plazo y esto lo hace extremadamente móvil. El único otro activo de envergadura que posee un banco —un personal talentoso— es también muy móvil. De ahí que un banco que sea incapaz de movilizar la ganancia se halle en peligro de perder con rapidez sus principales activos dejándolos en manos de sus competidores más atrevidos.

La combinación de una política monetaria carente de solidez con una laxitud regulatoria creó un ambiente en el que la competencia obligó a los bancos a asumir riesgos que, si bien se creía que eran de poca monta, podrían, en el caso de materializarse, conducir a la economía entera al abismo. Las prácticas de compensación también desempeñaron un papel en la asunción de riesgos por los banqueros. Un director ejecutivo o general protegido cómodamente contra la pérdida posee un incentivo para tomar altos riesgos a fin de maximizar el valor de sus opciones sobre valores. Además, entre mayor sea la compensación pagada al director ejecutivo mayor será su estímulo para maximizar las ganancias corporativas a corto plazo, ya que amasará una gran cantidad de dinero a corto plazo e, incluso si la empresa colapsa más adelante, él no tendrá que devolver el dinero a menos que gran parte de su compensación sea retrasada. A resultas de la debacle financiera, las corporaciones están empezando a retrasar la compensación a una escala mucho mayor que antes de que, al contemplar el espectáculo de la crisis, se hiciera trizas la fe del público en que el mercado de la retribución dada a los ejecutivos podía ser librado a sus propias fuerzas para que se autorregulara.

Dejando a un lado el punto de la compensación, si no se debiera culpar a los banqueros por tomar los riesgos que enfrentaron, ¿qué ocurre con los prestatarios? ¿No debería una persona tomar una decisión con más elementos que la de comprar una casa para la cual le podría alcanzar sólo si los precios de la vivienda siguieran al alza, de modo que a través de la casa adquiriría un capital accionario que le permitiera refinanciar con una hipoteca al alcance de su bolsillo? Se trataba de un juego de apuestas, es cierto, pero ¿era irracional? Supongamos que usted compra una casa por 200 000 dólares. No aporta dinero alguno (no tiene nada de dinero), y el trato es que no tendrá que pagar intereses durante dos años, pero a partir de inicios del tercer año la tasa de interés será de un rígido 8%... que usted no podrá cubrir. Si es de esperar que los precios de la vivienda continúen subiendo a una tasa de 10% por año (como lo estaban haciendo hasta que alcanzaron un pico en mayo de 2006), entonces dentro de dos años su casa tendrá un valor de 240 000 dólares, y tendrá la capacidad de refinanciarla con una hipoteca estándar de 30 años a

una tasa de interés a su alcance. Si los precios de la vivienda no se elevan, entonces en el tercer año abandonará la casa y volverá a rentar, y probablemente el único costo además de los costos de mudanza será una reducción en su solvencia crediticia. Las personas que hicieron esta apuesta con los ojos abiertos no eran irracionales, aunque no deja de ser una pregunta abierta cuántas de entre ellas tenían los ojos bien abiertos.

Cada vez que acontece un colapso súbito en los valores de los activos, se echa la culpa a la especulación. Ésta no es una reacción ridícula, porque las burbujas de precios de los activos son el resultado de la "especulación" de que los precios de alguna clase de activos seguirán subiendo el tiempo suficiente para que la compra de ese activo sea rentable. Lo que sí es ridículo, y resulta característico de las críticas endilgadas a la especulación, es culpar sólo a los corredores de bolsa que especulan sobre una caída en los precios de una clase de activos. Sin esos corredores, las burbujas alcanzarían un tamaño aún más abultado antes de estallar.

Es fácil apostar a un aumento en el precio futuro de algún activo simplemente comprando ese activo; es un reto mayor apostar a una caída en ese precio. Si es una acción (u otro valor, incluido un bono), uno puede tomarlo prestado y consentir en venderlo a alguien en una fecha especificada en el futuro a un precio especificado.[42] Si, tal como usted lo espera, el precio baja, puede comprar la acción que ha convenido en vender a un precio más bajo que el precio de venta, entregarle la acción que tomó en préstamo al comprador y recibir el pago al precio acordado, embolsarse la diferencia y hacer entrega de la acción barata que acaba de comprarle a la persona de quien tomó en préstamo la acción para la especulación, con lo que se redondea finalmente la transacción celebrada. Tal es la práctica conocida como "vender en corto", y es riesgosa, porque el precio de la acción puede elevarse en lugar de caer (así que tendrá que comprarla a un precio más alto que el precio especificado en el contrato de venta cuando llegue el tiempo de devolver una acción equivalente a la que tomó prestada), así como costosa, porque tiene que pagarle el interés a la persona de quien obtuvo el préstamo.

Como una alternativa a la venta en corto, puede comprar una permuta de cobertura por incumplimiento crediticio, con la que nos topamos antes en IV.6 considerándola como una forma de seguro sobre la deuda. Pero no es solamente eso. Es también un instrumento especulativo. Si hay un bono que no es de su propiedad pero al que espera acceder por incumplimiento (podría ser un bono respaldado por una colección de hipotecas), puede comprar una permuta de cobertura por incumplimiento crediticio que le pagará la pérdida en el valor del bono.

[42] ¿Tiene uno que pedir en préstamo la acción que vende en corto? Si no lo hace, ¿no podría vender en corto más acciones de las reservas de la compañía de las que fueran acciones pendientes? ¿Podría tal venta en corto "al desnudo" ser algo malo? Véase Sullivan & Long, Inc. *vs.* Scattered Corp., 47 F.3d, 857 (7th Cir., 1995).

A semejanza de otros especuladores, los vendedores en corto y los compradores de permutas de cobertura por incumplimiento crediticio que aseguran la deuda de desconocidos son impopulares porque están negociando en torno a una calamidad y, por consiguiente, abrigan esperanzas de que se presente. Pero culpar a los especuladores cuando el precio de un activo cae como resultado de la actividad especulativa es parecido a echarle la culpa a un termómetro de que baje la temperatura. A condición de que los especuladores no difundan falsos rumores acerca de los activos a cuyo descenso en el precio anhelan asistir, ni se involucren en otras clases de fraude, su actividad agrega elementos a la información que circula en el mercado y de esta manera tiende a traer consigo una alineación más completa y rápida entre los precios y los valores subyacentes.

Si en los inicios del decenio de 2000 hubiera sido rampante la venta en corto de bonos basados en hipotecas, el precio de éstos habría disminuido debido a que el alto nivel de la venta en corto habría constituido un indicio de que se habían generalizado las dudas sobre la posibilidad de que los precios de la vivienda siguieran elevándose. Cuando los precios de los bonos sufren una caída los rendimientos se elevan, ya que la tasa de interés pagada por un bono es un porcentaje fijo del valor aparente del bono. (Por ejemplo, si el valor de un bono que paga un interés de 2% decae hasta ser de la mitad, la tasa de interés para los compradores del bono se eleva a 4%.) Si las tasas de interés sobre los bonos hipotecarios se hubieran elevado y, por consiguiente, los precios de la vivienda hubieran bajado (por la razón de que el interés hipotecario es un costo bastante grande de la compra de una casa), la nación se podría haber ahorrado la burbuja de la vivienda cuya explosión hizo que se precipitara la crisis económica.

XVI.15. La Falacia del Nirvana y la economía darwiniana

Bien recuerdo haber escuchado, cuando llegué por primera vez a la Universidad de Chicago en 1969, la expresión "Falacia del Nirvana", utilizada para describir la creencia entonces predominante en la profesión de la economía de que las fallas en el mercado podrían y deberían rectificarse mediante la intervención del gobierno, la cual, según se suponía, era apolítica y estaba exenta de costo efectivo. Dicha creencia no tenía solidez; las fallas cometidas por el gobierno son un lugar común, en parte a causa del mundo de la política, en parte a causa de la dificultad intrínseca a muchas de las tareas cuya ejecución se le asigna al gobierno.

Los críticos de la Falacia del Nirvana no negaban la existencia de fallas en el mercado; solamente querían que los costos de esas fallas se compensaran con el costo de la intervención gubernamental antes de decidir si convenía recomendar esta última. No obstante, con el paso de los años, los economistas

conservadores (y, vista en conjunto, la economía es una profesión conservadora) optaron por ver a los mercados como el Nirvana —en tanto autorreguladores— y, de esta manera, por considerar a virtualmente cualquier forma de regulación gubernamental como superflua, costosa e ineficiente… incluida la regulación de la actividad bancaria. (La desregulación de la actividad de la banca había comenzado bajo la administración del presidente Carter y se completó bajo el segundo periodo del presidente Clinton.) Las crisis financieras, las depresiones, aun las recesiones severas pasaron a ser juzgadas como una cosa del pasado. Alan Greenspan, cuando fungía de presidente de la Junta de la Reserva Federal, se erigió en un vocero influyente de esta postura. Si bien esta complacencia era de índole bipartidista, se afianzó particularmente durante la administración del segundo presidente Bush, lo que acarreó las repercusiones seriamente adversas de que nos ocupamos en la sección precedente de este capítulo.

Semejante complacencia no ha acabado de esfumarse del todo. Los economistas dan cuenta de su propia incapacidad para haber previsto el colapso de la vivienda y, a la vuelta de la esquina, del colapso crediticio sosteniendo que, si fuera dado predecir las catástrofes, éstas serían evitadas; puesto que éstas suceden, la cuestión es que debe de ser imposible anticiparse a ellas. Ése es un argumento ridículo. La mayor parte de las catástrofes son previstas y como resultado podemos librarnos de ellas. Algunas no se pueden impedir. La pregunta, entonces, es por qué no. La respuesta en el caso presente la aporta la Falacia del Nirvana y la complacencia profesional que ésta engendra.

Un punto relacionado con este asunto es que aferrarse demasiado al modelo de elección racional, ignorando los aspectos psicológicos que lo complican (temas de la economía y las finanzas del comportamiento), desempeñaron un papel todo menos benigno en la dificultad que tuvieron los economistas para anticipar y responder al derrumbe financiero. El énfasis que John Maynard Keynes asignó en su magna obra *Teoría general del empleo, el interés y el dinero* (1936) a las motivaciones para el atesoramiento o acaparamiento, a los "espíritus animales" (como un acicate para la inversión, de cara a los efectos económicos de la incertidumbre irreductible (véase I.5)) y a la mayor resistencia que oponen los trabajadores a las reducciones en el salario nominal que a las habidas en el salario real es indispensable para comprender el ciclo económico, pero fue en gran medida ignorado en el periodo previo a nuestra crisis.

El Nirvana es una especie de cielo (el cielo como un estado de tranquilidad espiritual de carácter profundo y permanente). La naturaleza, de dientes y garras color sangre, es una especie de infierno. Aun así, el darwinismo proporciona un importante correctivo al pensamiento económico autocomplaciente que estoy describiendo. Las analogías entre la evolución biológica y la economía comercial son numerosas, como lo sugiere el hecho de que palabras como "competencia" y "equilibrio" se presenten abundantemente tanto en la biología como en la economía de tendencia evolucionista. Muchos resulta-

dos económicos habituales pueden modelarse como productos de una lucha por la sobrevivencia entre competidores, sin necesidad de suponer que haya de parte de ninguno de ellos una intención deliberada de maximizar la ganancia.

El establecimiento explícito de una analogía entre la economía y la teoría darwiniana alcanzó su apogeo en las décadas subsiguientes a la publicación en 1859 de *El origen de las especies*: en el movimiento que vino a ser conocido como el darwinismo social. En su forma extrema, el darwinismo social abogaba por la crianza eugenésica, con el objeto de mejorar a la raza humana, y por la eliminación del socorro a los menesterosos y otras políticas redistributivas, que, según se juzgaba, interferían con la lucha por la supervivencia, y, de modo concomitante, por la supervivencia de los más aptos. Con todo, los darwinistas sociales habían cometido un error grave, si bien común, que residía en confundir la condición de estar en buenas condiciones físicas y tener aptitudes con la cualidad de ser bondadoso. La "aptitud" y la condición óptima en la teoría darwiniana consisten en la adaptación al ambiente, no en la voluntad de mejorar las cosas… lo cual me retrotrae a las actividades bancarias y, en un sentido más general, a la competencia económica. Un ambiente competitivo favorece a las empresas que pueden adaptarse a tal ambiente; y, de esta forma, para determinar si un mercado está funcionando bien desde el punto de vista social en conjunto, uno tiene que tratar de entender el ambiente, así como la conducta que en el mundo comercial favorece mejor a una empresa para sobrevivir y prosperar en él. Un análisis que corresponde a esa motivación resulta una aplicación fructífera de la teoría darwiniana a los mercados competitivos.

Debido a que la evolución darwiniana marcha rumbo a la aptitud en vez de a la bondad, una industria de alto riesgo como la de la banca puede estar óptimamente adaptada a su ambiente económico aun si los riesgos que asume generan un serio peligro para la economía como un todo. Ésa es la argumentación a favor del seguro federal de depósito y otras regulaciones tradicionales (si bien hoy en día ya desmanteladas en la mayoría de los casos) de la solvencia bancaria.

Los economistas del Nirvana dan por sentado que los mercados "funcionan". En efecto, éstos funcionan, pero de la misma manera en que funciona la evolución biológica. La evolución biológica ha producido una maravillosa variedad de formas de vida, y la evolución económica ha producido una maravillosa variedad de productos y servicios que han incrementado en una gran proporción el bienestar humano. No obstante, ambos tipos de evolución producen "buenos" resultados sólo como productos colaterales de la lucha por la supervivencia. No siendo regulados, todo lo que pueden alcanzar es el logro de la aptitud. Eso puede ser bastante bueno en la mayoría de las industrias, pero no es lo suficientemente bueno, para la salud general de la economía, en el ámbito de la actividad bancaria.

XVI.16. ¿Son codiciosos los banqueros?

He dicho que los banqueros no debieran ser culpados por el derrumbe de 2008, puesto que simplemente estaban haciendo lo que tenían que hacer a fin de sobrevivir en el escenario del mercado; estaban tan libres de culpa por el desplome como los leopardos o los halcones u otros animales depredadores lo están por la crueldad que infligen a otros animales en su busca de comida. Una explicación más popular del derrumbe es que los banqueros son codiciosos. El significado económico de "codicia" es dar un peso mayor del promedio a la riqueza en la función de utilidad que uno maneja. La codicia se correlaciona con la preferencia por el riesgo porque en una sociedad competitiva resulta difícil obtener riqueza (de no ser por la herencia recibida) sin asumir riesgos. La banca moderna tal como la conocemos es muy arriesgada y, a diferencia de otras actividades económicamente riesgosas, tales como la actuación y la escritura de novelas y la interpretación de conciertos, ofrece una oportunidad más que minúscula de cosechar recompensas lucrativas.

Para repetir (y amplificar ligeramente) la discusión previa sobre la actividad bancaria en este capítulo, el capital del banco es a corto plazo y, por ello, puede desaparecer sin mayor o sin ninguna advertencia (una estampida bancaria), en tanto que sus activos (sus préstamos) son a largo plazo y pueden carecer de liquidez y, por consiguiente, ser difíciles de vender en el caso de que el banco requiriera remplazar parte de su capital. El seguro de depósitos reduce el riesgo de insolvencia que acecha al banco, pero el seguro de depósitos representa un costo para el banco, que por lo tanto tendrá un incentivo para aumentar su base de capital mediante un préstamo no asegurado. También puede decidir compensar el costo de sus reservas (el efectivo sobre el cual no recibe ningún rendimiento), así como amplificar el diferencial entre su costo del capital prestado y su rendimiento sobre la inversión, realizando inversiones más arriesgadas con sus fondos prestados que las hechas en préstamos hipotecarios, en bonos municipales o corporativos, en pagarés del Tesoro y en otros usos bancarios convencionales; en breve, puede decidirse a especular. Y, ya que el riesgo se materializará menos probablemente a corto plazo que a largo plazo (un 1% anual de riesgo de impago se vuelve formidable sólo cuando se proyecta a través de 10 o 20 años o aun un periodo más prolongado), un banquero que tenga una alta probabilidad de obtener ganancias muy cuantiosas durante los siguientes 10 años puede sentirse bien compensado por afrontar un pequeño riesgo de quiebra durante ese lapso. El capital financiero de un banco no es el único a corto plazo: también lo es su capital humano. Es muy escaso el capital humano financiero que parecería ser específico de la empresa, a juzgar por la velocidad a la que se desplazan los banqueros de empresa en empresa. Cualquiera de éstas que tenga un capital a corto plazo se halla sometida a una presión ingente para competir con fe-

rocidad, toda vez que se encuentra en constante peligro de perder su capital a manos de competidores más fieros y menos escrupulosos, quienes les pueden ofrecer a sus inversionistas y a sus empleados clave ingresos más altos.

Estas circunstancias hacen de una industria bancaria no regulada una selva darwiniana, en la que los banqueros son los depredadores y sus clientes (y cada uno de los ellos) son la presa. Ello explica por qué los banqueros son propensos a hacer recortes y jugar trucos: a tomar excesivos riesgos desde el punto de vista social considerado éste como algo distinto de su punto de vista privado. Así pues, la respuesta es sí: los banqueros pueden ser descritos como codiciosos. Pero no debiera entenderse esta palabra como una etiqueta peyorativa, pues las personas que están dispuestas a confrontar riesgos con la esperanza de obtener riqueza contribuyen de manera vital al crecimiento económico.

Lecturas recomendadas

Alexander, Janet Cooper, "The Value of Bad News in Securities Class Actions", 41 *UCLA L. Rev.,* 1421 (1994).

Bebchuk, Lucian A., y Holger Spamann, "Regulating Bankers' Pay", 98 *Geo. L. J.,* 247 (2010).

"Behavioral Economics and Investor Protection Conference" (trabajos presentados en la conferencia), 44 *Loyola U. Chi. L. Rev.,* 1323 (2013).

Brealey, Richard A., y Stewart C. Myers, *Principles of Corporate Finance* (7ª ed., 2002).

Easterbrook, Frank H., y Daniel R. Fischel, "Mandatory Disclosure and the Protection of Investors", 70 *Va. L. Rev.,* 669 (1984).

Fischel, Daniel R., y Sanford J. Grossman, "Customer Protection in Futures and Securities Markets", 4 *J. Futures Mkts.,* 273 (1984).

Gennaioli, Nicola, Andrei Shleifer y Robert W. Vishny, "A Model of Shadow Banking", 48 *J. Finance,* 1331 (2013).

Gordon, Jeffrey N., "The Puzzling Persistence of the Constrained Prudent Man Rule", *N. Y. U. L. Rev.,* 52 (1987).

Grossman, Sanford J., y Oliver D. Hart, "Corporate Financial Structure and Managerial Incentives", en *The Economics of Information and Uncertainty,* 107 (John J. McCall [comp.], 1982).

Keynes, John Maynard, *The General Theory of Employment, Interest and Money* (1936).

Kitch, Edmund W., "The Theory and Practice of Securities Disclosure", 61 *Brooklyn. L. Rev.,* 763 (1995).

La Porta, Rafael, Florencio Lopez-de-Silanes y Andrei Shleifer, "What Works in Securities Laws?", 61 *J. Fin.,* 1 (2006).

Langbein, John H., "The Uniform Prudent Investor Act and the Future of Trust Investing", 81 *Iowa L. Rev.,* 641 (1996).

———, y Richard A. Posner, "Market Funds and Trust-Investment Law", partes I y II, 1976 *Am. Bar Foundation Research J.,* 1 (1977).
———, y Richard A. Posner, "Social Investing and the Law of Trusts", 79 *Mich. L. Rev.,* 72 (1980).
Posner, Richard A., *A Failure of Capitalism: The Crisis of '08 and the Descent into Depression* (2009).
———, *The Crisis of Democratic Capitalism* (2010).
Rajan, Raghuram, *Fault Lines: How Hidden Fractures Still Threaten the World Economy* (2010).
Shleifer, Andrei, *Inefficient Markets: An Introduction to Behavioral Finance* (2000).
Taylor, John B., *Getting Off the Track: How Government Actions and Interventions Caused, Prolonged, and Worsened the Financial Crisis* (2009).
The Squam Lake Report: Fixing the Financial System (2010).

Problemas

1. Evalúe las aseveraciones siguientes: "Debido a su ponderación, la cartera [S&P] del índice [es decir, la cartera del mercado] tenderá siempre a desempeñarse mejor [que los fondos mutuos convencionales] en un mercado declinante, pero no tan bien en un mercado creciente. Dado que los periodos de precios crecientes tienden a exceder a los periodos de precios declinantes, un fondo de índice es una fórmula para un perdedor sólido, consistente a largo plazo. ¿Quién necesita eso? ¿Por qué es tan prudente?"[43] "Parecería una cuestión de prudencia casi elemental reducir la volatilidad de una cartera de acciones a niveles altos del mercado y aumentarla a niveles bajos."[44]

2. ¿Debiera permitirse a un administrador fiduciario (en ausencia de lenguaje explícito en la escritura constitutiva del fideicomiso) que invirtiera en acciones inmobiliarias? ¿En hipotecas inmobiliarias? ¿Debiera permitírsele operar un negocio con los fondos fiduciarios, por oposición a una inversión meramente pasiva?

3. Supóngase que un instrumento fiduciario estipula que la viuda del fideicomitente habrá de recibir el ingreso derivado de los activos de la disposición mientras viva, y una vez fallecida éstos pasarán a sus hijos. Para cumplir con su deber de imparcialidad ante los distintos beneficiarios y no anular el valor de la fundación, ¿debería el administrador fiduciario invertir los activos de ésta en acciones comunes y canalizar los dividendos hacia la viuda? ¿Debería invertirlos en bonos y darle los intereses? ¿O debería hacer algo di-

[43] Roger F. Murray, "Investment Risk in Pension Funds: The Pension Benefit Guaranty Corporation View", en *Evolving Concepts of Prudence: The Changing Responsibilities of the Investment Fiduciary in the Age of ERISA,* 37, 39 (Financial Analysts Research Foundation, 1976).

[44] *Ibid.,* 40.

ferente? Véase Alyssa A. DiRusso y Kathleen M. Sablone, "Statutory Techniques for Balancing the Financial Interests of Trust Beneficiaries", 39 *U. San Francisco L. Rev.,* 261 (2005).

4. La regla de la prioridad absoluta de la ley federal de quiebras, según recordamos (véase antes XV.5), está diseñada para asegurar que las reclamaciones de los acreedores más antiguos se satisfagan plenamente antes de que se permita a los acreedores y a los accionistas recibir ningún dinero de la corporación. Supongamos que una corporación se está reorganizando en la quiebra, habiendo dejado de pagar sus bonos, por los que debe ahora 100 dólares de principal más los intereses vencidos. Si la corporación se liquida de inmediato, sólo obtendrá 50 dólares por la venta de sus activos. Pero si continúa operando como una empresa en marcha, podrá hacerlo por un año más, al final del cual se habrán depreciado todos sus activos, con el cuadro XVI.1 de ganancias posibles de las operaciones del año:

CUADRO XVI.1

Ganancias ($)	*Probabilidad*
0	1/3
110	1/3
220	1/3

La cifra de las ganancias esperadas —el promedio ponderado de todos los resultados de desenlaces posibles— es de 110 dólares. Suponiendo que una tasa de descuento de 10% es apropiada para el nivel de riesgo asociado a los resultados posibles de la corporación, su valor presente como un negocio con buen funcionamiento se obtiene descontando en 10% el resultado esperado de 110 dólares al final de un año, de modo que es igual a 110/1.10, o sea, 100 dólares. ¿Cumpliría con el espíritu de la regla de la prioridad absoluta un plan de reorganización que otorgara a los tenedores de bonos nuevos bonos que les dieran derecho a recibir un máximo de 110 dólares a partir de las ganancias disponibles al final de las operaciones del año, y a los accionistas nuevas acciones comunes que les dieran derecho a cualquier residuo después de que se hubieran pagado los bonos nuevos? ¿O debieran entregarse a los tenedores de bonos todas las acciones comunes de la corporación?

5. La "negociación de súper alta frecuencia" (una modalidad de lo que se denomina "negociación cuantitativa" o "inversión cualitativa") remite al empleo de técnicas matemáticas y estadísticas avanzadas y computarizadas cuyo objetivo es detectar las pautas de cambio acelerado en los precios de las acciones; estos datos se convierten luego en la base para efectuar negociaciones fundamentalmente instantáneas en el curso de las cuales determinadas

acciones a menudo adoptan posiciones por sólo unas fracciones de segundo. El resultado de ello es un mercado de valores más eficiente. Discuta los beneficios sociales y los costos sociales del incremento en la eficiencia del mercado que trae consigo la negociación de súper alta frecuencia.

Quinta Parte

EL DERECHO Y LA DISTRIBUCIÓN DEL INGRESO Y LA RIQUEZA

XVII. LAS DESIGUALDADES DEL INGRESO, LA JUSTICIA DISTRIBUTIVA Y LA POBREZA[1]

XVII.1. La medición de la desigualdad

El ingreso monetario se encuentra distribuido desigualmente. En la gráfica XVII.1 el porcentaje del ingreso recibido, acumulado desde el más bajo hasta el más alto, se expresa como una función del porcentaje de las unidades familiares, acumulado desde el más bajo hasta el más alto. Si el ingreso se distribuyera igualmente entre todas las unidades familiares del país, la función sería la línea recta llamada *línea de la igualdad*. En cada punto de esa línea, la fracción del ingreso recibida por un porcentaje dado de las unidades familiares es exactamente igual a ese porcentaje: 20% de las unidades tienen 20% del ingreso, 55% de las unidades tienen 55% del ingreso, etc. Entre más inclinada sea la distribución efectiva, menos igualitaria será. En la actualidad, 20% de los hogares más pobres en los Estados Unidos reciben aproximadamente 5% del total del ingreso personal total de la nación, 20% de los más ricos reciben alrededor de 50%,[2] y 1% de los más ricos reciben 20%.[3] Entre las naciones ricas del mundo, los Estados Unidos, Gran Bretaña e Italia tienen la distribución del ingreso más inequitativa, mientras que Dinamarca, Finlandia, los Países Bajos, Noruega, Suecia y Alemania tienen la más equitativa.[4] La desigualdad ha crecido en algunos países acaudalados, en especial en los Estados Unidos, pero ha disminuido en otros; no existe una tendencia general que sea visible.[5] Volveremos al tema del incremento en la desigualdad del ingreso en los Estados Unidos en XVII.3, después.

Las estadísticas sobre la desigualdad del ingreso no proveen una guía clara para la política social, ni una imagen clara de la desigualdad del ingreso. Al tomar una fotografía instantánea de los ingresos de un año, tales estadísticas comparan engañosamente a personas que se encuentran en etapas diferentes del ciclo vital (una distorsión que se corrige sólo parcialmente al no contar a los niños, la mayoría de los cuales no tienen ingresos). Por ejemplo, las estadísticas colocarán a un abogado joven que acaba de unirse a una

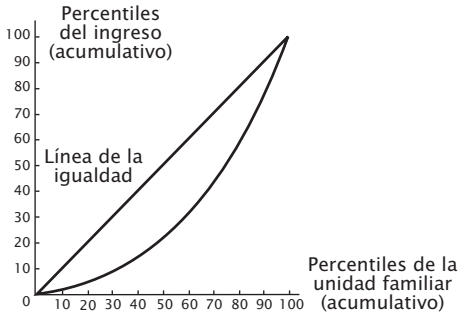
[1] Véase Facundo Alvaredo *et al.*, "The Top 1 Percent in International and Historical Perspective", 27 *J. Econ. Perspectives*, 3, verano de 2013.

[2] Timothy M. Smeeding, "Public Policy, Economic Inequality, and Poverty: The United States in Comparative Perspective", 86 *Soc. Sci. Q.,* 955, 967 (2005) (cuadro 1).

[3] Alvaredo *et al.*, "The Top 1 Percent…", p. 3.

[4] Smeeding, antes nota 2, en 959.

[5] *Ibid.,* en 958 (Fig. 1).



GRÁFICA XVII.1

firma legal y a un socio antiguo de la misma empresa en dos clases de ingresos diferentes, aunque ambos podrían ganar las mismas cantidades a lo largo de sus vidas (en realidad, el abogado joven probablemente ganará más).

Además, limitadas al ingreso pecuniario, las estadísticas dejan de lado muchos factores que son muy importantes para el bienestar económico —aun definido en un sentido muy estrecho—, pero difíciles de cuantificar. Comparemos a dos familias: en una de ellas trabajan ambos esposos y cada uno de ellos gana 40 000 dólares anuales; en la otra, sólo trabaja el esposo y gana 80 000 dólares. El ingreso pecuniario de las familias es el mismo, pero el ingreso real de la segunda es mayor. La esposa se queda en casa porque sus servicios valen más en el hogar, para la familia, que el ingreso que obtendría de un empleo externo. Si ella es inteligente, dinámica y bien instruida, su ingreso alternativo podría ser muy elevado: igual o mayor que el de su marido (véase antes V.1). Es precisamente porque las mismas cualidades pueden convertirla en un ama de casa y madre especialmente competente que su valor en el hogar podría exceder a su valor en el mercado.

Aun en los empleos remunerados, una parte importante de la paga es a menudo no pecuniaria. Los maestros, por ejemplo, reciben una gran parte de su ingreso bajo la forma de largas vacaciones. Las estadísticas no sólo dejan fuera cierto ingreso no pecuniario, sino que también cuentan como ingreso pecuniario algunos ingresos que son, en efecto, rembolsos de costos. Sabemos que quienes desempeñan empleos peligrosos o desagradables recibirán, en igualdad de otras condiciones, salarios mayores que quienes desempeñan empleos sin estas características indeseables; pero el ingreso real de los dos grupos podría ser el mismo. Es también importante el riesgo en el sentido de la incertidumbre económica, ya no del peligro. Supongamos que el ingreso de un inventor exitoso es de 500 000 dólares al año, mientras que el de un inventor fracasado es de 0, y que la probabilidad de que un inventor triunfe es de 1 en 10 (omitimos las probabilidades intermedias). El ingreso esperado de los inventores es entonces de 50 000 dólares, y es igual al de los empleados públicos que ganan con seguridad 50 000 dólares cada año. Si se fijara un tope

de 50 000 dólares a los ingresos de los inventores exitosos, nadie escogería una carrera de inventor; el ingreso esperado de tal carrera sería de sólo 5 000 dólares anuales. En efecto, 90% del ingreso aparentemente elevado del inventor exitoso de nuestro ejemplo es una compensación por el riesgo de la pérdida. No es sólo que el inventor afronte un riesgo anormal; el empleado público recibe una parte de su ingreso bajo la forma de una seguridad del empleo que no es gravable. Los jueces federales tienen una seguridad del empleo mayor aún que la de los empleados públicos, y además reciben pensiones extraordinariamente generosas que no pagan impuestos; el valor presente neto de los ingresos esperados de un juez federal de 58 años de edad superará probablemente al de los abogados privados que tienen ingresos anuales sustancialmente mayores.

Consideremos una distribución simple de cuatro ingresos: el de un carpintero de 20 años de edad que gana 20 000 dólares; el de un estudiante universitario de 20 años de edad que no gana nada; el de un carpintero de 30 años de edad que gana 30 000 dólares, y el de un licenciado de 30 años de edad, un contador, que gana 40 000 dólares. Aparentemente, hay una desigualdad considerable, pero es posible que en efecto no haya ninguna. El ingreso de 0 del estudiante representa una inversión en educación, la que recupera a lo largo de su vida de trabajo bajo la forma de un sueldo mayor. Los 10 000 dólares adicionales que gana a la edad de 30 años, por comparación con el ingreso de un carpintero que ya estaba trabajando cuando el estudiante se encontraba todavía en la escuela, podrían representar simplemente el pago con intereses de una parte de la aportación de capital con la forma de colegiaturas y un ingreso sacrificado que él o su familia realizaron cuando él era un estudiante años atrás.[6]

Los gobiernos federal y estatales, y en menor medida otras instituciones (hospitales, fundaciones, etc.), proveen diversos beneficios, tales como educación, protección policiaca, atención médica, pensiones, subsidios para los pobres e instalaciones recreativas, sin ningún cobro directo. Estos beneficios tienen un impacto diferencial sobre el bienestar de distintos grupos de ingresos. Es importante que no sólo se mida este impacto, sino que se determine también cómo se distribuyen los costos (primordialmente los impuestos) entre las clases de ingresos. Por ejemplo, si la carga de la educación pública de los niños pobres recae primordialmente en las familias pobres, la educación

[6] Pero, en la medida en que la educación fue pagada por el Estado, la familia, donantes de la universidad o cualquier otro aparte del estudiante, el "pago" del capital se describe mejor como la recepción de servicios de un donativo que como una forma de la amortización. Por lo que toca a la importancia empírica de la inversión educativa en la explicación de las diferencias de ingresos, véanse por ejemplo Michael Hout, "Social and Economic Returns to College Education in the United States", 38 *Ann. Rev. of Sociology*, 379 (2012); Larry L. Leslie y Paul T. Brinkman, *The Economic Value of Higher Education* (1988); Jacob Mincer, *Schooling, Experience, and Earnings* (1974).

pública no podrá lograr ninguna transferencia neta de riqueza desde las clases más ricas de la sociedad. Pero, en efecto, los pobres, aunque pagan impuestos mayores de lo que podría pensarse (sobre todo indirectamente), reciben en transferencias más de lo que pagan en impuestos,[7] aunque quizá sólo un poco más.[8]

Un factor que milita contra la igualdad real es el hecho de que el prestigio (fama, estatus social, respeto, etc.), una forma del ingreso no pecuniario que es importante para muchas personas, en nuestra sociedad se correlaciona de ordinario en forma positiva con el ingreso. Es cierto que algunas ocupaciones relativamente mal pagadas, como las del clero y los militares, son bastante prestigiosas. Pero la posesión de un ingreso grande o una riqueza sustancial genera por lo menos una deferencia ostensible y a menudo un respeto real, y estas cosas aumentan el valor de la riqueza para el poseedor. Y quienes se encuentran en el fondo de la distribución del ingreso reciben a veces un respeto negativo bajo la forma de un tratamiento rudo o insensible por parte de la policía, el personal médico, los arrendadores y los vendedores de tiendas de abarrotes o servicios, lo que deprime más aún su ingreso real.

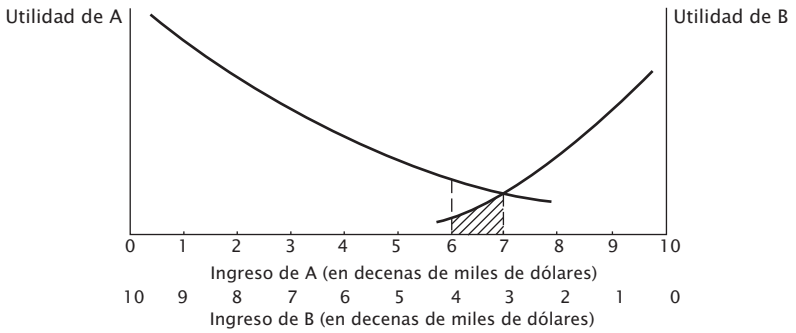
Otra complicación reside en el grado en que la preferencia por el riesgo y la aversión al riesgo, así como las tasas de descuento personal, inciden sobre el ingreso real considerado independientemente de la cantidad de dólares que uno gane. La aversión al riesgo ocasionaría que el ingreso esperado de 20 000 dólares del inventor fuera menor al ingreso seguro de 20 000 dólares del servidor público; de tener una preferencia por el riesgo, el ingreso del inventor sería más elevado. Una tasa de descuento personal alta hará que las ganancias futuras sean de un valor mucho menos sustancial para un trabajador que sus ganancias en la actualidad, lo que significa un revés potencial para el inventor en el supuesto de que no pueda esperar un éxito inmediato.

XVII.2. ¿LA DESIGUALDAD ES INEFICIENTE?

¿Aumentaría el bienestar económico la adopción de políticas destinadas a que la sociedad se aproximara a la igualdad? El principio de la utilidad marginal decreciente del dinero podría tentarnos a concluir que una transferencia de una persona más rica a una persona menos rica tenderá a aumentar la suma de las utilidades de las dos personas: una pérdida de un dólar perjudica al millonario menos de lo que la ganancia de un dólar agrada al indigente. Esto

[7] Richard A. Musgrave y Peggy B. Musgrave, *Public Finance in Theory and Practice,* 264 (5ª ed., 1989) (cuadro 12-4).

[8] Benjamin I. Page, *Who Gets What from Government,* 135-144 (1983). Lo que es más, los Estados Unidos dedican una porción de su PIB mucho más pequeña que otras naciones ricas a los esfuerzos por reducir la pobreza. Timothy Smeeding, "Poor People in Rich Nations: The United States in Comparative Perspective", 20 *J. Econ. Perspectives*, invierno de 2006, en 69.



GRÁFICA XVII.2

es cierto incuestionablemente, pero generalizar a partir de este ejemplo extremo y concluir que una redistribución sustancial de la riqueza, de las personas de ingresos mayores a las personas de ingresos menores, aumentará seguramente la utilidad total sería peligroso aun bajo la premisa poco realista de que la redistribución no cuesta nada.

En la gráfica XVII.2 se representa la utilidad marginal en el eje vertical y el ingreso monetario en el eje horizontal. La curva de la izquierda muestra la utilidad marginal de A como una función de su ingreso monetario. Su ingreso actual es de 70 000 dólares y el área bajo la curva a la izquierda de 70 000 es entonces su utilidad total. La curva tiene una pendiente negativa, lo que implica una utilidad marginal decreciente. Es decir, si el ingreso de A aumentó de 70 000 a 80 000 dólares, el aumento de su utilidad (el área bajo la curva entre estas dos cifras) sería menor que si su ingreso hubiese sido de 60 000 y aumentara a 70 000 dólares. El ingreso de B es de 30 000 dólares. Su utilidad total es el área bajo la curva en el lado derecho del diagrama, a la derecha de 30 000 dólares en su escala. Si se transfieren 10 000 dólares de ingreso de A a B, la utilidad de A disminuirá en el área bajo su curva de utilidad marginal entre 70 000 y 60 000 dólares en su escala, mientras que la utilidad de B aumentará en el área bajo su curva de utilidad marginal entre los mismos dos puntos (30 000 y 40 000 dólares en su escala). La última área, sombreada en el diagrama, es menor que el área que mide la disminución de la utilidad de A como resultado de la transferencia. La transferencia ha generado una disminución de la utilidad total. La razón es que, mientras que A y B experimentan una utilidad marginal decreciente, A obtiene de 1 dólar una utilidad mayor que la de B en la región relevante.

Dado que no se conocen la forma ni la altura de las curvas de utilidad marginal de las personas, no puede descartarse la posibilidad de que estas curvas de las personas más ricas sean mayores, en promedio, que las de las más pobres. Se ha sugerido que el supuesto más plausible, dada esta ignorancia, es que las curvas de utilidad marginal son las mismas entre todos los grupos de

ingresos; según esta concepción, sería probable que la igualación de los ingresos aumentara la utilidad total.[9] ¿Pero no sería posible que el ingreso y su utilidad marginal correspondiente estuvieran correlacionados *positivamente*, por lo menos dentro de ciertos límites, según la teoría de que las personas que trabajan arduamente para ganar dinero y logran hacerlo son, en promedio, quienes valúan más el dinero, habiendo renunciado a otras cosas tales como el ocio para tenerlo? Además, el hecho de que, a pesar de la utilidad marginal decreciente del ingreso, los pobres sean compradores ávidos de billetes de lotería implica que esperan que al volverse ricos se elevará su curva de utilidad marginal del ingreso, puesto que la ganancia pecuniaria esperada de jugar la lotería es negativa porque una parte de las ganancias se utiliza para subsidiar la operación de la lotería. Y hemos omitido los costos de la redistribución del ingreso. Si tales costos fueran sustanciales, como veremos más adelante, podrían eliminar las ganancias de la utilidad total generadas por la redistribución.[10]

Dejando de lado los costos de la redistribución, podemos poner un poco de orden en el análisis comparando dos conjuntos de razones de las diferencias de los ingresos: diferencias en el gusto por el ingreso en relación con otros bienes (llamémosles "ocio") y diferencias en las oportunidades o la habilidad para ganar dinero. Si sólo hubiera diferencias de gustos, los ricos *deberían* tener mayores utilidades marginales del ingreso que los pobres; de otro modo, los ricos demandarían el mismo ocio, y sus ingresos bajarían. Si las diferencias en el gusto por el dinero carecen de importancia y las diferencias del ingreso se generan por la suerte, la salud, los sesos, etc., esto implica que la utilidad marginal del ingreso no se correlaciona con la riqueza y, por lo tanto, que si la redistribución de los ricos a los pobres no costara nada, aumentaría la utilidad agregada. Aun así, no aumentaría la *riqueza* de la sociedad. La riqueza permanecería sin cambio o, más realistamente, cuando se toman en cuenta los costos de la redistribución, se reduciría a resultas de la redistribución, lo que en consecuencia no podría defenderse por razones de la eficiencia, si ésta se define como se sugirió en el capítulo I.

¿Cómo cambiarán las cosas si hay mucha envidia en la sociedad? La envidia para un economista ilustra el fenómeno de las utilidades negativas interdependientes: si A envidia a B, una disminución de la utilidad de B aumentará

[9] Abba Lerner, *The Economics of Control,* 35-36 (1944).

[10] Estas consideraciones se omiten con frecuencia. Por ejemplo, C. Edwin Baker, "Utility and Rights: Two Justifications for State Action Increasing Equality", 84 *Yale L. J.,* 39 (1974), tras admitir (véase *ibid.* en 41-42, n. 12) que el análisis de Lerner, aun si se aceptara, sólo establecería una probabilidad de que la redistribución aumentara la utilidad total, olvida la calificación y afirma que si se quita al rico para darle al pobre aumentará seguramente la utilidad total (*ibid.* en 45). También se olvida el reconocimiento que hace el autor de que los costos de la redistribución deben restarse de cualesquiera ganancias (*ibid.* en 48, n. 26), lo que le permite concluir —sin ninguna referencia nueva a los posibles costos de la redistribución— que ésa es una meta económicamente justificada de la política económica (*ibid.* en 55-56, 58).

la utilidad de A. (La envidia es así lo contrario del altruismo, el caso de las utilidades positivas interdependientes.) La existencia de la envidia podría parecer un poderoso apoyo normativo para la redistribución, a menos que la envidia pueda llamarse de algún modo una preferencia "mala". Un economista no se siente inclinado a criticar preferencias que no generan externalidades negativas significativas, y en realidad la envidia podría generar externalidades positivas al incitar a los individuos a ser más productivos a fin de que su estatus se acerque más al de los envidiados. Pero, aun si la envidia no tiene nada de malo, no provee de un apoyo robusto para la redistribución. La envidia depende de la capacidad para ligar nuestros sentimientos con los de otros que no son como nosotros, sentir sus alegrías como nuestros dolores y sus dolores como nuestras alegrías. Esto implica que la envidia podría ser más aguda cuando las diferencias del ingreso son pequeñas que cuando son grandes, porque resulta más fácil sentir "empatía", positiva o negativa, por las personas que son como nosotros. La observación famosa de Tocqueville en el sentido de que "la mayor igualdad tiende a generar comparaciones envidiosas: a medida que se vuelven más iguales, a los individuos les resulta cada vez más difícil soportar su desigualdad".[11]

La igualdad del ingreso podría ser mucho menos importante que el ingreso total para mantener la paz social y la estabilidad política. En una sociedad donde la mayor parte de la población es pobre y sólo hay una pequeña clase rica, la mayoría no tiene ningún interés en el régimen democrático, de modo que no lo defenderán contra un golpe de Estado. Siendo pobres, es probable que sientan (aunque a menudo erróneamente) que no tienen nada que perder de un cambio en el sistema de gobierno. Compárese esto con una sociedad que, independientemente de que tenga o no una clase baja numerosa, está integrada primordialmente por personas ricas: si, además, como parece plausible, la libertad política es un bien superior, se la exigirá en mayor proporción si el ingreso per cápita es alto. Por añadidura, las personas acaudaladas a un tiempo tienen menos probabilidades de ser rebeldes y son más difíciles de reprimir.[12]

Podría parecer que, si en una democracia hay gran desigualdad del ingreso, el votante mediano tendrá un mayor incentivo para apoyar los impuestos fuertemente progresivos, ya que serán grandes las oportunidades existentes para la redistribución de los ricos a quienes no son ricos; y, como veremos en el siguiente capítulo, los impuestos fuertemente progresivos reducen la eficiencia económica. Por lo tanto, si los ingresos pudieran volverse más igualitarios con medidas menos ineficientes que la tributación altamente progresiva, el ingreso nacional podría ser mayor *porque* estaba distribuido más igualitariamente. Pero este análisis depende decisivamente de la existencia y la viabili-

[11] Véase Raymond Boudon, "The Logic of Relative Frustration", en *Rationality and Revolution,* 245 (Michael Taylor [comp.], 1988).

[12] John F. O. Bilson, "Civil Liberty-An Econometric Investigation", 35 *Kyklos,* 94, 103 (1982).

dad política de medidas de redistribución más eficientes que la tributación progresiva y de la validez de un modelo de resultados políticos del votante mediano: un modelo que deja de lado el papel de los grupos de interés en el proceso político (véase el capítulo XX).

XVII.3. El aumento de la desigualdad: causas y consecuencias

En las décadas recientes la desigualdad en el ingreso en los Estados Unidos ha aumentado con celeridad, y a una tasa acelerada. En 1982, 1% en el extremo superior de los estadunidenses que generaban ingresos recibían 12.8% del ingreso total; para 1997, esta cifra se había incrementado a 17%, y en los siguientes tres años saltó a 21.3%. Desde entonces se ha estabilizado y, como se mencionó en la primera sección de este capítulo, es ahora de 20% (aunque, en realidad, la estimación más reciente es de 22%). La tendencia de una desigualdad en aumento en el largo plazo parece deberse principalmente a que la gente bien educada obtiene retornos cada vez mayores a consecuencia del viraje ocurrido en la economía hacia métodos de producción y distribución que requieren la aplicación de conocimiento e inteligencia más que de fuerza o vigor físico, pero asimismo a otros factores, como una caída en el tipo marginal del impuesto sobre la renta, o tipo marginal del impuesto sobre la renta de las personas físicas (IRPF).[13] Otros factores incluyen las ganancias sin precedentes de la industria financiera, recibidas en su mayor parte por ejecutivos financieros bajo la forma de ingresos muy elevados, así como el agudo incremento en primas para el seguro de salud suministrado por el empleador: el aumento fue de casi 80% entre 2000 y 2009. Gran parte de esto constituye un pago nominal realizado por el empleador, pero, dado que se trata de un costo de la mano de obra, sustituye a los incrementos salariales y, de esta manera, contribuye a un estancamiento del salario por debajo del nivel de quienes reciben una paga elevada. Entre 1997 y 2008, el ingreso medio en los hogares estadunidenses cayó 4% después del ajuste por la inflación, y entre 2007 y 2011 cayó 8.9%. Una media no es un promedio; el ingreso promedio se elevó porque se elevaron los ingresos de quienes reciben sueldos más altos, y, de este modo, el efecto consistió en incrementar la desigualdad del ingreso.

No existen razones para creer que estas tendencias no habrán de seguir dándose; y, mientras el desempleo no descienda a un nivel normal, es difícil encontrar algo que pudiera servir para prevalecer sobre las tendencias si éstas persisten.

Al sopesar los efectos de la desigualdad en el ingreso, es importante distinguir entre ingreso monetario y nivel de vida. En la misma proporción en que la calidad de los bienes y servicios aumente (en gran medida a conse-

[13] Véase Alvaredo *et al.*, XVI.1 antes, n. 3, en 18.

cuencia de la innovación tecnológica entendida en un sentido amplio que incluye los nuevos métodos en las actividades comerciales, así como el progreso en el campo de la ciencia y los inventos) más rápido que sus costos, el nivel de vida aumentará aun cuando no ocurra lo mismo con los ingresos. La calidad de los servicios de salud sigue aumentando, y una parte del incremento en las primas de los seguros de salud corresponde al pago por ese aumento de calidad. El costo ajustado a la calidad de la electrónica para el consumidor se ha hundido en el mismo periodo.

Aun así, por más que haya aumentado el nivel de vida para la mayoría de las personas cuyos ingresos no se han elevado o incluso han llegado a descender, esto no aliviaría los problemas políticos que con toda probabilidad genera un incremento en la desigualdad económica. Las personas dan por consabidas complacientemente la mayoría de las mejoras en bienes y servicios, en lugar de considerarlas una compensación de un ingreso monetario en declive. Por otro lado, asimismo, las restricciones en liquidez pueden excluir a las personas del acceso a muchas de las mejoras en cuestión; éste es un problema para aquellos que no pueden cubrir un seguro de salud.

La ansiedad económica puede acarrear consecuencias económicas desastrosas al aumentar la demanda de proteccionismo comercial, restricciones a la inmigración, protecciones a los sindicatos, otra serie de medidas anticompetitivas y, en fin, subsidios gubernamentales; puede igualmente crear un resentimiento de clase y, así, abrir el paso a medidas regulatorias ineficientes. De esta manera, las medidas adoptadas para reducir la desigualdad económica, en especial las encaminadas a elevar el ingreso medio de los hogares (las cuales se diferencian de las que persiguen el mismo propósito disminuyendo el nivel de ingresos de los privilegiados, que pueden tener efectos serios de desincentivación), podrían incrementar la eficiencia económica reduciendo las presiones políticas en pos de políticas ineficientes. Tal era la lógica que justificaba que naciones capitalistas adoptaran "medidas socialistas", empezando con los programas de Bismarck de un seguro de salud, contra accidentes y discapacidad y, asimismo, de pensiones para la vejez proporcionados por el gobierno, todos ellos diseñados para servir al capitalismo de coraza contra el comunismo y otras ideologías políticas radicales. ¡Y su programa funcionó!

Esto sucedió en ese entonces; ahora la red de seguridad social es demasiado onerosa para ser expandida sin poner en jaque la solvencia de los Estados Unidos (y de muchos otros países), en vista de una deuda pública muy elevada y en crecimiento constante (véase XVII.10 después). Las únicas medidas que afrontarían el estancamiento monetario sin aumentar nuestra deuda pública serían subsidios que pudieran tener una justificación realista en su condición de inversiones rentables en capital humano (tales como la subvención pública de la colegiatura en las universidades), al igual que cambios regulatorios virtualmente exentos de costo como el de eliminar el subsidio fiscal para los beneficios de salud provistos por el empleador, eliminar o al menos

reducir muchos otros subsidios fiscales, instituir pruebas para verificar si las personas tienen derecho a recibir beneficios de Medicare y seguridad social, atenuar el efecto de ciertas leyes de seguridad y ambientales a fin de reducir costos para las empresas, restarles fuerza a los sindicatos de maestros y a otros sindicatos de empleados públicos, ahondar en la privatización de los servicios públicos, reducir las barreras arancelarias y dar cabida a una mayor inmigración de trabajadores altamente calificados.

El economista Raghuram Rajam especula que el aumento de la desigualdad es una causa indirecta del colapso financiero de 2008 y la crisis económica que le siguió.[14] Para mantener el nivel de consumo de cara a la declinación en el ingreso del hogar, la Reserva Federal, como sabemos, presionó las tasas de interés a corto plazo y, al proceder de ese modo, así como al propiciar las hipotecas de tasas ajustables, contribuyó a que subieran los precios de la vivienda y a la vez se facilitara el préstamo otorgado al consumidor. Los propietarios de viviendas podían sacar préstamos casa-capital accionario con base en el crecimiento de un capital accionario derivado de la casa y utilizar los beneficios para financiar el consumo, así como incurrir en una mayor deuda con la tarjeta de crédito y en el ámbito no relacionado con la vivienda. La deuda remplazó al consumo, con lo que se preservaron las ilusiones acerca de una prosperidad ininterrumpida y siempre en expansión.

XVII.4. La teoría contractual de la justicia distributiva y su aplicación a la seguridad social

El filósofo John Rawls ha sostenido[15] que la distribución del ingreso y la riqueza es justa sólo si ninguna otra distribución mejorara la situación de los miembros de la sociedad que se encuentran en la peor situación. La distribución justa podría ser muy desigual si, por ejemplo, el impacto negativo de una distribución más igualitaria sobre el incentivo para trabajar fuese tan sustancial que la tajada mayor que recibirían quienes estuviesen en peor situación fuese menor en términos absolutos que la tajada relativamente menor que recibirían bajo una distribución menos igualitaria.

El hecho de que se consideren primordiales los intereses de quienes se encuentran en la peor situación parecería violar el principio aceptado por la virtual totalidad de los economistas modernos en el sentido de que son arbitrarias las comparaciones interpersonales de la utilidad (felicidad). La justificación de Rawls es que resulta plausible suponer que, si pudiéramos encuestar de algún modo a todos los individuos que han integrado alguna vez

[14] Véase Raghuram Rajan, *Fault Lines: How Hidden Fractures Still Threaten the World Economy* (2010).

[15] En su famoso libro *A Theory of Justice* (1971) [edición en español: *Teoría de la justicia*, trad. de María Dolores González, FCE, Madrid-México, 1995].

la sociedad, o que la integrarán alguna vez, en sus posiciones originales (es decir, antes de que sepan cuál será su lugar en la sociedad), se observaría que tales individuos prefieren un conjunto de arreglos que maximicen la posición de quienes se encuentren en la peor situación. Esto supone que todos los individuos sienten aversión por el riesgo, lo que está muy bien, pero es menos plausible suponer que todos los individuos sienten una aversión enorme por el riesgo.

Comparemos dos conjuntos de arreglos sociales, X y Y. X conduce a una distribución del ingreso en la que el ingreso medio del 10% más pobre de los miembros de la sociedad es de 10 000 dólares anuales, mientras que el ingreso medio del total de los miembros de la sociedad es también de 10 000 dólares. Y conduce a una distribución del ingreso en la que el ingreso medio del 10% más pobre es de 9 000 dólares, pero el ingreso medio del total de los miembros de la sociedad es de 40 000 dólares. Es improbable que escojan X quienes se encontraban en la posición original, aunque sientan aversión por el riesgo.[16] Pero esto sólo demuestra que no se establece la forma específica del principio de la justicia de Rawls (maximin). Su enfoque básico no se ve afectado, y rompe ingeniosamente con el estancamiento que surgió en la sección anterior cuando tratamos de comparar las utilidades de los individuos *después* de que se sacaran los billetes de la lotería de la vida. La utilidad para un individuo podrá reducirse cuando se le quite su billete ganador y se le entregue a alguien más, aunque la lotería no se haya realizado justamente. Si A parte un pastel y le entrega a B la parte más grande, B podría resistirse a darle una porción de ella a C, quien recibió la parte más pequeña. Pero su resistencia es irrelevante para la cuestión de si el método de partición del pastel fue justo.

Rawls nos pide hacer lo que hemos hecho muchas veces en este libro: imaginar el contenido de un contrato que no puede celebrarse en el mercado debido a los elevados costos de transacción. Quienes se encuentran en la posición original saben que la riqueza de la sociedad podría dividirse en muchas formas. Si sienten aversión por el riesgo, presumiblemente desearán alguna protección contra la obtención de una rebanada muy pequeña (a menos que el pastel sea muy grande) o nada en absoluto. El principio de Rawls les otorga una protección excesiva, pero esto puede corregirse fácilmente (en efecto

[16] Rawls descarta esta posibilidad alegando primordialmente que los grandes aumentos del bienestar de quienes se encuentran en mejor situación aumentarán así seguramente también el bienestar de quienes se encuentran en la peor situación, por lo menos cuando se toman en cuenta otras salvaguardias incorporadas en su principio de la justicia (tal como la competencia abierta por las posiciones). Rawls, nota 1, en 157-158. Esto omite la posibilidad de que las reducciones sustanciales del bienestar de quienes se encuentran en la mejor situación pudieran conducir a un mejoramiento trivial del bienestar de quienes se encuentran en la peor situación, en cuyo caso podría requerirse un conjunto de arreglos institucionales que afectaran muy fuertemente a los ricos de acuerdo con el principio de la justicia de Rawls.

se corrigió… ¡mucho tiempo antes de que escribiera Rawls![17]). Dada una probabilidad igual de que lo sea cualquiera, cada persona en la posición original deseará maximizar la utilidad esperada de su boleto en la lotería de la vida, y estas expectativas se maximizan maximizando la utilidad total. Dado que la aversión al riesgo afecta a la utilidad, las políticas sociales maximizadoras de la utilidad (dependiendo del costo) incluirán algunas provisiones redistributivas: algún seguro social o una "red de seguridad" para a aquellos a quienes no les sonría la fortuna.

Por desgracia, la teoría de la justicia distributiva de Rawls no tiene ningún contenido operativo. Aparte del problema de determinar cuánta aversión al riesgo habrá de suponerse, tenemos el problema de decidir quién será contado como en peor situación, que Rawls reconoce pero no trata de resolver. Si quien se encuentra en peor situación es una sola persona, se considerarán injustas algunas medidas que serían justas si quienes se encuentran en peor situación fuesen, por ejemplo, los miembros del decil de ingresos más bajos. Es también decisivo que el universo relevante sea una sola sociedad o todo el mundo, ya que en el primer caso las medidas tomadas para mejorar la suerte de los estadunidenses en la peor situación a costa de los extranjeros serían prohibidas por el principio de justicia de Rawls. Es también crucial juzgar si la peor situación deba entenderse estrictamente en términos del ingreso monetario, con todos los problemas que esa medida implica, o en términos más amplios. Existe igualmente el problema de que no está claro cuáles políticas particulares generan cuáles resultados. El resultado de estas incertidumbres es que la teoría de la justicia de Rawls parece enteramente compatible, por una parte, con el socialismo total y, por la otra parte, con el capitalismo del *laissez-faire*, dependiendo de cuánta aversión al riesgo deba suponerse que tienen las personas situadas en la posición original, cuán estrechamente se defina el grupo de los colocados en la peor situación en cuyo interés deben operar todas las políticas, y cuán eficaces se consideren las instituciones públicas en relación con el mercado libre.

La teoría del contrato social a la manera de Rawls puede utilizarse para cuestionar la creencia convencional de que el programa de seguridad social representa una redistribución masiva de la riqueza de los jóvenes a los viejos. Es cierto que la seguridad social es un programa de pago inmediato y no un plan de pensiones autofinanciado. Los beneficios de la seguridad social se pagan con los impuestos de quienes están empleados ahora, antes que con los ahorros de los receptores. Pero el sistema puede considerarse como un contrato justo entre los jóvenes y los viejos. Los adultos en edad de trabajar pagan impuestos para apoyar al sistema de la escuela pública, en efecto prestando dinero a los jóvenes para que compren capital humano. Los jóve-

[17] John C. Harsanyi, "Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking", 61 *J. Pol. Econ*., 434 (1953). Véase también Kenneth J. Arrow, "Some Ordinalist-Utilitarian Notes on Rawls's Theory of Justice", 70 *J. Philo.,* 245 (1973).

nes pagan el préstamo cuando se vuelven adultos en edad de trabajar y la generación que pagó por su educación en la escuela pública llega a la edad de la jubilación. Lo hacen sufragando, a través de los impuestos de la seguridad social, algunos de los gastos de subsistencia y médicos de la generación de sus padres.[18] Este punto se pasa por alto en los análisis convencionales de los efectos redistributivos porque tales análisis sólo consideran los impuestos de la seguridad social, no los impuestos escolares, al calcular la medida en la que los receptores de los beneficios de la seguridad social pagaron por ellos cuando estaban empleados.

Un argumento relacionado es que, contra las apariencias, el sistema de seguridad social está confiriendo beneficios a la generación actual en edad de trabajar, al igual que a los receptores de la seguridad social. En ausencia de la seguridad social, la carga del mantenimiento de los viejos recaería en gran medida en los miembros más jóvenes de sus familias. Por supuesto, hablar de la "carga" de obligaciones asumidas voluntariamente, con toda la renuencia que se quiera, es alejarse de las convenciones habituales del análisis económico. Pero si bien muchos hijos adultos son suficientemente altruistas hacia sus padres ancianos para estar dispuestos a incurrir en costos sustanciales en términos de dinero, tiempo, irritación, aflicción y aun repulsión, antes que descuidarlos o abandonarlos, preferirían grandemente echar sobre otros hombros la carga del cuidado de sus padres, por lo menos en parte. El cuidado de los ancianos difiere así de los tipos de consumo de los que los individuos obtienen placer. La seguridad social reparte el componente monetario de la carga del cuidado de los ancianos entre toda la generación actual. Al hacerlo así, podría producir una equiparación mejor de los costos y los beneficios de los servicios gubernamentales. El mayor beneficio de la seguridad social al aliviar la carga del cuidado de los ancianos se confiere a las familias pequeñas; un hijo único afronta una carga mayor de sostener a sus padres en la ancianidad que un hijo que tenga muchos hermanos con los cuales compartir esa carga. Pero el mayor beneficio de la seguridad social para las familias pequeñas se balancea por el hecho de que son las familias grandes las que obtienen un beneficio mayor de los impuestos escolares, porque tienen más hijos que educar y los impuestos escolares no varían con el número de los hijos del contribuyente.

XVII.5. La movilidad social

Es importante distinguir entre desigualdad e inmovilidad: entre diferencias en el ingreso dadas en un momento determinado y cambios en el ingreso re-

[18] Gary S. Becker, *A Treatise on the Family,* 369-374 (edición ampliada, 1991). El principio de nuestro plan de pensiones públicas en las que éstas se pagan con las cuotas (impuestos) que dan los que están empleados actualmente debe distinguirse de su diseño efectivo, en el que se han combinado consideraciones prácticas y políticas para producir muchas anomalías.

lativo a través del tiempo; o, en otras palabras, entre desigualdad estática y desigualdad dinámica. Si el coeficiente intelectual o CI fuese puramente genético y, por ello, solamente heredado, si el ingreso fuese una función lineal del CI (en lo que tal vez se está convirtiendo, con los veloces avances en la tecnología que presagian un futuro en el que los robots operados por brillantes ingenieros en computación se encargarán de todo el trabajo manual), los ingresos de los hijos como adultos serían muy similares a los de sus padres y la movilidad intergeneracional sería escasa, no sin hacer, empero, la salvedad de que la posición social no es puramente una función del ingreso monetario. Los científicos, los compositores de música clásica, los poetas, los líderes políticos y otros funcionarios, los oficiales militares y los sacerdotes constituyen ejemplos de personas que tienen un estatus social elevado pero salarios mediocres. Y adviértase que la correlación entre la posición social o el ingreso de los padres y los hijos se sostendría para cualquier rasgo puramente genético que fuese valuado, como el atletismo o el talento musical, y no sólo para el CI. Comoquiera que fuese, aunque todavía existiera la movilidad intergeneracional en el ingreso monetario, el ingreso en su sentido más cabal de utilidad no sería muy móvil. Cierto, el CI no es puramente genético (por lo general se cree que es mitad genético y mitad el resultado de otras condiciones innatas, por ejemplo, las que acompañaron al embarazo y el nacimiento, además de las relativas al ambiente temprano y a la salud física y mental), pero de todos modos tiene un componente genético que es sustancial. Y la creciente diversidad en el apareamiento (los semejantes junto a los semejantes), atribuible en parte a un descenso en la discriminación y en parte a la mayor búsqueda de compañeros maritales que trae consigo la Internet, probablemente significará un mayor aumento en la variación del CI y, así, conducirá a un aumento de las ganancias para el estrato de los poseedores del CI más alto en el futuro cercano.

De acuerdo con este análisis, un nivel bajo de movilidad intergeneracional es congruente con la meritocracia, pero los Estados Unidos son una meritocracia imperfecta, en parte a causa de la pauta de inversión, tanto privada como pública, en las carreras y las perspectivas de vida de los hijos. Los padres acaudalados invierten mucho en la educación de sus hijos contratando tutores, pagando las muy elevadas cuotas cobradas por las escuelas, colegios y universidades (tanto públicos como privados), otorgando generosos donativos a esas instituciones y financiando actividades extracurriculares que causan gran impresión en los representantes de la oficina de admisiones de esos centros de estudio. Los colegios y universidades compiten en pos de hijos adinerados, a quienes estiman por verlos en el futuro como exalumnos convertidos en generosos patrocinadores. Los padres utilizan contactos personales a fin de conseguir buenos trabajos para sus hijos; en ocasiones los contratan para la empresa familiar; y, por supuesto, les dan, prestan y legan dinero. Debido a que las escuelas públicas son financiadas

principalmente a través de tributaciones a las propiedades locales, las mejores escuelas públicas tienden a ubicarse en áreas residenciales de grandes ingresos, y, por consiguiente, los chicos más ricos que van a las escuelas públicas en vez de acudir a las privadas tienden a ir a las mejores escuelas públicas. De hecho, si no hay buenas escuelas públicas en el barrio donde vive una familia acaudalada, la familia enviará a sus hijos a una escuela privada.

No resulta accidental que entre los países ricos el nivel de movilidad intergeneracional sea más bajo en los Estados Unidos que en los cuatro países escandinavos (que poseen el nivel de movilidad intergeneracional más alto) y en Canadá, Australia, Nueva Zelanda, Alemania, Japón, Francia y España, y que sea ligeramente más elevado en los Estados Unidos que en Italia y el Reino Unido. Estas clasificaciones están altamente correlacionadas con la desigualdad del ingreso (los Estados Unidos ocupan la segunda posición superior en desigualdad del ingreso, después del Reino Unido, en los países recién enlistados). Las familias provistas de una gran fortuna que financian a sus hijos con un acopio de generosidad crean más familias acaudaladas en la siguiente generación.

Con la finalidad de maximizar la productividad, uno encontraría deseable que el gobierno identificara a los hijos poseedores de un CI más alto en las familias pobres o de clase media baja y les proporcionara educación y servicios ligados a la educación equivalentes a los que esas familias adineradas brindan a sus hijos, usualmente dueños de un CI alto. Esta necesidad se presenta con agudeza en los Estados Unidos porque nuestro nivel de movilidad intergeneracional, en relación con los de otros países, es bastante bajo. No es mucho el esfuerzo que se está realizando para satisfacer esta necesidad. Las razones estriban probablemente en que los hijos no votan (cabría sostener que a cada progenitor debería concedérsele un voto extra por cada uno de los hijos o hijas que vivan con él), y en que los padres de clase baja y media baja tienen un escaso ascendiente político en relación con las personas de la tercera edad, las adineradas, los pensionistas del servicio civil y otros grupos influyentes en la esfera política.

XVII.6. Los costos de la pobreza y las limitaciones de la caridad privada

El extremo en la desigualdad del ingreso que llamamos pobreza es un concepto relativo. En Etiopía o Bangladesh, pobreza significa algo diferente de la pobreza en los Estados Unidos. Y, dentro de la misma cultura, el concepto cambia a lo largo del tiempo. Comparando estimaciones de diversos periodos en nuestra historia del ingreso mínimo necesario para ubicar a una familia de tamaño determinado por encima del nivel de la pobreza, vemos que ese nivel se ha elevado sostenidamente, aun tras corregir por la disminución del

valor del dólar por efecto de la inflación.[19] (En otras palabras, la pobreza tiene un componente relativo, en contraposición a la medida de la privación simplemente absoluta.[20]) Actualmente, 15% de los estadunidenses tienen ingresos inferiores al nivel de la pobreza… el mismo porcentaje que en 1982.[21]

La pobreza impone, a quienes no son pobres, ciertos costos que ameritan, por razones estrictamente económicas (es decir, de maximización de la riqueza), que se incurra en algunos costos para reducirla. La pobreza en medio de la abundancia aumenta la tasa de la delincuencia.[22] El ingreso sacrificado de una ocupación alternativa legítima es bajo para alguien que tiene escasa capacidad de obtener ingresos en ocupaciones legítimas, mientras que la proximidad de la riqueza aumenta el rendimiento esperado del delito. El sector delictivo es en efecto el empleador de última instancia.

El costo de la desutilidad que tiene la pobreza también se impone a los altruistas ricos. El altruismo es mucho más fuerte entre los miembros de la familia que entre extraños,[23] pero aun un débil sentimiento de altruismo podría generar una transferencia voluntaria de ingreso si hay grandes disparidades en el ingreso. Si A es en extremo rico y B en extremo pobre, la sustracción de 1 dólar del ingreso de A podría disminuir la utilidad de A mucho menos que la cantidad en que agregar 1 dólar al ingreso de B aumentaría la utilidad de B: supongamos que 100 veces menos. Entonces A obtendría un beneficio neto

[19] Oscar Ornati, "The Poverty Bank and the Count of the Poor", en *Inequality and Poverty,* 167 (Edward C. Budd [comp.], 1967).

[20] Véanse, por ejemplo, Shaohua Chen y Martin Ravallion, "More Relatively-Poor People in a Less Absolutely-Poor World", 59 *Rev. Income & Wealth*, 1 (2013); Martin Ravallion y Shaohua Chen, "Weakly Relative Poverty", 93 *Rev. Econ. & Stat.*, 1251 (2011).

[21] Esto excluye, sin embargo, los beneficios no consistentes en efectivo, y se ha estimado que si esos beneficios se incluyeran en los ingresos de las personas que figuran allí el número de personas consideradas como pobres podría descender por lo menos 42% (esto es, de 15 a 9%). U.S. Dept. of Commerce, Bureau of the Census, "Characteristics of Households and Persons Receiving Selected Noncash Benefits: 1982", en 5 (Current Population Reports, Consumer Income, serie P-60, núm. 143, enero de 1984). Un estudio de la Oficina del Censo halló que la tasa de pobreza había disminuido a 12.5% en 2003 y habría caído a alrededor de 9.5% si los beneficios no concedidos en efectivo (incluidos Medicaid y Medicare) se hubieran incorporado en el cálculo del ingreso. U.S. Dept. of Commerce, Bureau of the Census, "Alternative Poverty Estimates in the United States: 2003", en 5 (cuadro 1) (junio de 2005).

[22] Isaac Ehrlich, "Participation in Illegitimate Activities: An Economic Analysis", en *Essays in the Economics of Crime and Punishment,* 68, 94 (Gary S. Becker y William M. Landes [comps.], 1974); véase también B. Curtis Eaton y William D. White, "The Distribution of Wealth and the Efficiency of Institutions", 29 *Economic Inquiry,* 336 (1991). Es cierto que las tasas delictivas son frecuentemente bajas en los países pobres, aunque la desigualdad de la riqueza es a menudo mucho mayor en esos países que en los países ricos. Si la riqueza se encuentra altamente concentrada, los costos de su protección contra los delincuentes pueden ser bajos. Es cuando la riqueza se encuentra distribuida más ampliamente que los delincuentes tienen abundantes blancos atractivos. Véase antes la discusión de VII.2, acerca de si las desigualdades de la riqueza promueven el delito.

[23] ¡Por cuya razón se excluye correctamente a los hijos de padres que no son pobres del conteo de los pobres!

al entregar 1 dólar a B, siempre que A valuara el bienestar de B derivado de la donación en más que una centésima parte de la propia utilidad que él (A) derivaría de guardarse el dinero.

Esta observación nos ayuda a entender por qué la pobreza es un concepto relativo. Lo que resulta decisivo desde el punto de vista del altruismo débil que sentimos hacia los extraños no es el nivel del ingreso en la parte más baja de la distribución del ingreso, sino la brecha existente entre su ingreso y el del resto de nosotros.

Podría creerse que, dado que los ricos ganan con la reducción de la pobreza, el alivio de la pobreza puede dejarse en manos de la caridad privada; pero el altruista afronta un problema de gorroneo. En nuestro ejemplo, A obtendrá bienestar del aumento del ingreso de B, independientemente de que A sea o no la fuente del aumento. Naturalmente, a A le gustaría comprar este aumento de su bienestar al precio más bajo posible, de modo que tendrá un incentivo para retraerse en el otorgamiento de caridades con la esperanza de que otros lo hagan. Podría pensarse que, independientemente de lo que otros donen, la contribución de A aumentará la cantidad total de recursos destinados a un fin que él aprecia. Pero esto no es seguro. Su contribución podría llevar a otros a reducir sus aportaciones, porque ahora una aportación menor de parte de ellos comprará (en combinación con la aportación de A) la misma reducción de la pobreza. Por lo tanto, A obtendrá menos que 1 dólar de beneficios por cada dólar que aporte, y esto lo llevará a reducir su aportación.

Siempre que haya problemas de gorroneo habrá un argumento económico a favor de la intervención gubernamental: aquí un argumento a favor de que se obligue a la gente a contribuir al alivio de la pobreza, de modo que no pueda aprovecharse gratuitamente de las donaciones privadas para la caridad. Pero el argumento no es concluyente; deben considerarse los costos de la intervención. Estos costos son grandes por dos razones: *1)* el gobierno carece en general de la disciplina de la competencia y del incentivo de la maximización de las ganancias, y *2)* resulta socavado por la política. Las tareas asignadas al gobierno por la sociedad tienden a ser inherentemente difíciles. El programa por discapacidad de la seguridad social, por ejemplo, ha creado considerables desincentivos para el trabajo, aunque la discapacidad para trabajar sea un requisito formal para la recepción de los beneficios de discapacidad.[24] Se ha descubierto que los programas de bienestar, tales como el de la Ayuda para Familias con Hijos Dependientes (AFDC), ahora abolido, tie-

[24] Hilary Williamson Hoynes y Diane Whitmore Schanzenbach, "Work Incentives and the Food Stamp Program", 96 *J. Pub. Econ.,* 151 (2012); Ariel Fiszbein y Norbert Schady, "Conditional Cash Transfers: Reducing Present and Future Poverty", A World Bank Policy Research Report, 118 (2009); Robert Moffitt, "Incentive Effects of the U.S. Welfare System", 30 *J. Econ. Lit.,* 1 (1992); Howard P. Marvel, "An Economic Analysis of the Operation of Social Security Disability Insurance", 18 *J. Human Resources*, 393 (1982).

nen efectos negativos sorprendentemente grandes sobre la participación en la fuerza de trabajo: la participación de las madres en el caso de la AFDC.[25] Este programa también tenía una característica de ciclo vicioso: inducía un aumento del número de nacimientos fuera de matrimonio al disminuir el costo de esos hijos para sus madres[26] y así aumentaba el número de las personas elegibles para beneficios y, por ende, el costo del programa. En total se ha estimado que de cada dólar transferido por el gobierno en la AFDC se disipan 23 centavos.[27] Hallazgos como éstos dieron cierto apoyo a la gran reforma del bienestar federal mediante la promulgación en 1996 de la Ley de Responsabilidad Personal y Oportunidad de Trabajar, que abolió a la AFDC y la sustituyó por donativos en bloque otorgados a los estados bajo la condición de que se adoptaran criterios estrictos para la beneficencia que trataran de alentar el trabajo y desalentar la dependencia. La reducción de las listas de la beneficencia y el aumento del empleo de los receptores de beneficencia, que la Ley (en unión de la fuerte economía del decenio de 1990) parece haber generado,[28] demuestran que los pobres responden racionalmente a los incentivos financieros.

Otro aspecto negativo de las transferencias gubernamentales es su efecto sobre la caridad privada. Las transferencias gubernamentales son un sustituto de la caridad privada, de modo que debieran conducir a una disminución de tal caridad: un efecto incrementado por el hecho de que el impuesto necesario para sufragar el costo de la transferencia disminuye el ingreso disponible después del impuesto para que el contribuyente practique la caridad privada. No obstante, aunque fue detectado en un estudio temprano,[29] existen ahora dudas de que sea significativo el efecto de disminución de tal caridad.[30]

Hay otro argumento a favor de la red de seguridad. Una persona rica que sienta aversión por el riesgo deseará asegurarse contra la posibilidad de volverse pobre alguna vez en el futuro, debido a los malos negocios, la mala sa-

[25] Véase Moffitt, nota 24 antes, en 15-18.

[26] Véase, por ejemplo, Margaret F. Brinig y F. H. Buckley, "The Price of Virtue", 98 *Public Choice,* 111 (1999).

[27] Sheldon Danziger y Robert Haveman, "How Income Transfer Programs Affect Work, Savings, and the Income Distribution: A Critical Review", 19 *J. Econ. Lit.*, 975, 1020 (1981). Algunos de estos costos podrán deberse al diseño deficiente de los programas. Véase la sección siguiente.

[28] Véanse, por ejemplo, Gregory Acs & Pamela J. Loprest, *Leaving Welfare: Employment and Well-Being of Families That Left Welfare in the Post-Entitlement Era* (2004); Rebecca M. Blank, "Declining Caseloads/Increased Work: What Can We Conclude About the Effects of Welfare Reform?", Federal Reserve Bank of New York Economic Policy Review, septiembre de 2001, 25.

[29] Burton A. Abrams y Mark D. Schmitz, "The 'Crowding Out' Effect of Governmental Transfers on Private Charitable Contributions", 33 *Pub. Choice,* 30, (1978); véase después en XVIII.9 (discusión de las deducciones por caridad en el impuesto federal sobre el ingreso [renta]).

[30] Véase Daniel Tinkelman, "Revenue Interactions: Crowding Out, Crowding In, or Neither?", en *Handbook of Research on Nonprofit Economics and Management,* 18, 20-21 (Bruce Alan Seaman y Dennis R. Young [comps.], 2010).

lud, los cambios del mercado de trabajo y otras desgracias. Algunas de las causas de la pobreza pueden asegurarse en el mercado de seguros privado, pero un seguro amplio contra la pobreza es difícil de obtener, por dos razones: el azar o riesgo moral y la selección adversa. Quienquiera que comprara un "seguro contra la pobreza" tendría un incentivo menor para trabajar y ahorrar, y un incentivo mayor para tener hijos y tener un comportamiento económico riesgoso. Asimismo, el grupo de asegurados estaría dominado por personas que esperaran ser pobres, y por ello las primas de seguros se harían estratosféricas. El gobierno tiene facultades de las que carece una compañía de seguros privada para resolver estos problemas, sobre todo la que le permite alterar las condiciones básicas que afectan a los incentivos del asegurado. Por ejemplo, reduciendo las tasas del impuesto al ingreso para los pobres, puede aumentar el costo de sacrificar el trabajo en favor de la beneficencia. Quizá lo más importante de todo, "estar en la beneficencia" lleva consigo un estigma, lo que no ocurre con el hecho de recoger el valor de una póliza de seguros que usted pagó; este estigma disminuye el azar o riesgo moral del seguro social. Y el seguro social resuelve el problema de la selección adversa porque a nadie se le permite retirarse del grupo de los asegurados.

XVII.7. Transferencias ilimitadas en efectivo *versus* beneficios en especie

No se entienden bien las causas y los remedios de la pobreza. La concepción más simple es que la pobreza deriva de una falta de ingreso y se trata más eficazmente mediante pagos en efectivo incondicionales a los pobres. Este enfoque (impuesto negativo al ingreso) ha contado desde hace mucho tiempo con el apoyo de muchos economistas, y desde 1975 ha formado parte del código tributario federal como el crédito al impuesto al ingreso ganado, que provee créditos fiscales para los trabajadores de salarios bajos. El programa se expandió durante el decenio de 1990 y ahora provee cerca de 62 000 millones de dólares en créditos a más de 27 millones de familias.

Tres argumentos económicos se esgrimen contra el enfoque del impuesto negativo al ingreso:

1. Contra nuestro supuesto anterior, es posible que los ricos no se preocupen por la utilidad subjetiva de los pobres sino por el patrón de consumo de los pobres. Podrían desear que los pobres cuenten con un nivel decente de vivienda, nutrición y oportunidades educativas, aunque los mismos pobres prefirieran otras cosas, tales como ropas finas, automóviles caros y billetes de lotería. Los ricos podrían esperar que, si puede inducirse a los pobres a comprar algunas cosas en lugar de otras, la pobreza disminuirá a largo plazo; que si, por ejemplo, una familia pobre tiene un nivel decente de vivienda, alimentación y educación, es me-

nos probable que la generación siguiente sea pobre. Viendo las cosas en esta forma, el donante rico obtiene más por su dinero etiquetando su contribución para usos particulares por parte del receptor pobre.

2. Algunas personas son pobres porque son incompetentes en la administración del dinero; en tal caso, los donativos ilimitados en efectivo podrían ser despilfarrados, sin aliviar la pobreza.
3. Las transferencias ilimitadas en efectivo pueden causar desempleo.[31] Si, por ejemplo, una de cada cuatro familias tuviera un ingreso mínimo garantizado de 15 000 dólares, el jefe de tal familia no tendría ningún incentivo para tomar un empleo aunque se le pagara más que eso. Un empleo que pagara 310 dólares a la semana aumentaría su ingreso bruto en sólo cerca de 10 dólares a la semana, y su ingreso neto bajaría efectivamente. Trabajar es caro. El trabajador debe pagar por el transporte, por las ropas de trabajo, el impuesto al ingreso federal y estatal, etc., aun cuando el trabajo mismo no genere ninguna desutilidad (como podría ocurrir, por supuesto, en especial si el trabajo es sucio, arduo o peligroso), de modo que no hay ningún costo de oportunidad por la renuncia al ocio del estado de desempleado.

Una escala graduada de transferencias en efectivo conservaría algún incentivo para el trabajo. Supongamos, por ejemplo, que por cada dólar ganado la contribución gubernamental al ingreso familiar disminuye sólo 50 centavos, aunque el resultado fuese que el ingreso total de una familia, ganado tanto como subsidiado, excediera al ingreso mínimo garantizado de 15 000 dólares. Pero incluso un tipo marginal del impuesto sobre la renta de 50%, sin duda mejor que una de 100%, es todavía muy pesada y tendrá efectos de desincentivo. Una persona que ganara 15 000 dólares al año retendría sólo 7 500, lo que después de la deducción de los costos del trabajo sería una minucia. Y quienes no están en la beneficencia que estuvieran trabajando por salarios bajos podrían verse inducidos a abandonar la fuerza de trabajo. Esto aumentaría el costo del programa y disminuiría la oferta de mano de obra en los empleos mal pagados. Aunque la disminución de la demanda de tales empleos conduciría a un aumento de los salarios ofrecidos, y esto contrarrestaría en parte el efecto de desincentivo volviendo al trabajo más remunerativo, también aumentaría el costo y, por ende, el precio de los bienes y servicios dependientes de tal mano de obra, de un modo muy semejante al de un impuesto indirecto. Como veremos en el capítulo siguiente, los impuestos específicos son generalmente regresivos.

[31] Véanse Oliver Bargain y Karina Doorley, "Caught in the Trap? Welfare's Disincentive and the Labor Supply of Single Men", 95 *J. Pub. Econ*., 1096 (2011); Edgar K. Browning, "Effects of the Earned Income Tax Credit on Income and Welfare", 48 *Nat'l Tax J.,* 23 (1995); Philip K. Robins, "A Comparison of the Labor Supply Findings from the Four Negative Income Tax Experiments", 20 *J. Human Resources,* 567 (1985).

La tasa marginal del impuesto sobre la renta puede variarse a voluntad para mitigar los efectos de desincentivo que acabamos de discutir, pero una disminución sustancial de la tasa aumentaría grandemente el costo del programa. Si la tasa marginal del impuesto fuese de 25%, una familia de cuatro miembros donde el jefe de la familia ganara 12 000 dólares anuales recibiría todavía del gobierno un complemento al ingreso anual de 2 000 dólares; y tendría que ganar 20 000 dólares antes de dejar de ser elegible para cualquier complemento.[32]

Sin embargo, antes de condenar el enfoque del "impuesto negativo al ingreso" por sus efectos de desincentivo del trabajo, adviértase que los subsidios etiquetados tienen los mismos efectos. Por ejemplo, en virtud de que el derecho a recibir un complemento de la renta está condicionado a que el ingreso no pase de cierto nivel, el ingreso ganado por encima de ese nivel se grava en efecto por la cantidad del complemento del ingreso recibido.

Dejando de lado los efectos de desincentivo del trabajo, el problema más grave de nuestro sistema gubernamental (en realidad se trata de sistemas) de ayuda a los pobres en la actualidad, que, aparte del crédito fiscal al ingreso percibido, es un sistema de subsidios etiquetados y de beneficios en especie —vivienda pública, complementos para la renta, estampillas para alimentos, servicios legales y médicos gratuitos, capacitación para el trabajo, etc.—, es un problema de información.[33] La disposición de un consumidor, incluido un consumidor pobre, a pagar el precio de mercado de un bien en lugar de otro es una prueba del valor relativo de los bienes para él. No importa cómo haya obtenido los dólares que está usando para expresar su preferencia. Pero cuando el precio de un bien es cero resulta muy difícil estimar el valor del bien para el "comprador". Aun si los funcionarios de la pobreza tuvieran una idea general de la forma como los pobres ordenan los diversos bienes y servicios, esto permitiría sólo un juicio promedio que, dadas las diferentes necesidades de familias pobres particulares, conduciría a una superfluidad en muchos casos y a una insuficiencia en muchos otros. Un programa menos extenso de transferencias en efectivo podría generar un bienestar neto mayor para los pobres al mismo tiempo que disminuyera los costos para el contribuyente. Pero el problema de la información no tiene un solo aspecto, por supuesto. Como se mencionó antes, muchos pobres deben de carecer de buena información acerca de cómo salir de la pobreza.

El desempeño efectivo de nuestro sistema de transferencias etiquetadas deja mucho que desear. Un ejemplo particularmente interesante para los abogados es el programa por el que el gobierno, a través de la Corporación de Servicios Legales, provee asistencia legal gratuita a los pobres en materia civil. Este método de ayuda a los pobres podría impedir efectivamente que

[32] $5 000 – (0.25 × $12 000) = $2 000; $5 000 – (0.25 × $20 000) = 0.

[33] Ralph K. Winter, Jr., "Poverty, Economic Equality, and the Equal Protection Clause", 1972 *S. Ct. Rev.,* 41.

muchos pobres alcanzaran su patrón de consumo más eficiente. Dado que los fondos gubernamentales asignados a los servicios legales para los pobres no están disponibles para otros programas de ayuda a los pobres, el costo para la persona pobre de tener derecho a 100 dólares en servicios legales podría ser el beneficio que habría obtenido de la recepción de 100 dólares de algún otro bien o servicio o en efectivo. Muchos pobres podrían ser capaces de arreglárselas bien sin un abogado, ya sea porque tienen la fortuna de no padecer problemas legales o porque son suficientemente astutos para afrontarlos sin la ayuda de un abogado. Pero, dado que el abogado es gratuito,[34] lo usarán a menos que el valor de sus servicios supere al valor (a menudo escaso) de su tiempo para tratar con él. Enfrentado a una demanda excesiva de su tiempo, el abogado tratará de limitar sus servicios a aquellos cuyas necesidades de servicios legales parezcan más agudas, y, dado que esto requiere un juicio difícil, habrá inevitablemente muchos casos en los que un pobre recibe servicios legales que cuestan 100 dólares pero que sólo valen para él 50 dólares, o no recibe servicios legales que serían mucho más valiosos que su costo. Aparte de los problemas de información y de la competencia de los consumidores, si se entregaran a los pobres 100 dólares en efectivo en lugar de un abogado gratuito, éstos usarían el efectivo para contratar a un abogado en lugar de comprar alimentos, medicinas, educación o vivienda sólo cuando el valor de los servicios legales para ellos fuese de por lo menos 100 dólares.[35]

La disponibilidad de seguros legales es una cuestión relevante. Si el seguro está disponible, resulta fácil traducir un recibo de pago en efectivo en protección contra la eventualidad de no contar con un abogado, siempre y cuando, por supuesto, el seguro no sea demasiado caro. Si el seguro legal no está disponible, entonces cabe argumentar que proporcionar un abogado gratuito al indigente deviene una forma eficiente de seguro social.

Además, los servicios legales se utilizan típicamente en una disputa —en el caso de un pobre, con un arrendador, un cónyuge, un comerciante, una dependencia de la beneficencia, una compañía financiera, etc.—. Los esfuerzos legales que se hagan en nombre de una de las partes de la disputa aumentarán los costos para la otra, quien deberá aumentar *sus* esfuerzos legales o abandonar en provecho de la otra parte lo que esté en juego en la disputa. Estos costos son de ordinario costos marginales —una función de la producción—, de modo que, por un análisis que ya nos resulta familiar, se trasladarán (por lo menos en parte) a los clientes de la empresa. Si la mayoría de esos clientes son otros pobres, esto significa que los costos generados por el uso de un abogado por parte de un inquilino, cliente o deudor pobres

[34] No completamente, empero, si el tiempo del cliente posee un valor que se pierde al dialogar con el abogado.

[35] La asistencia legal para un pobre acusado de un delito podría plantear un problema especial. Véase después XXVI.2.

serán soportados en gran parte por otros pobres.[36] Si la disputa es con una dependencia gubernamental, el gasto legal adicional para la agencia será soportado indirectamente por los pobres, porque el costo incrementado del programa podría conducir a una disminución del monto o la cobertura de los beneficios.

Contra esto se puede sostener que es menos probable que el arrendador, el comerciante o el prestamista traten de defraudar o explotar de otro modo a los consumidores de ingresos bajos si saben que tienen acceso a servicios legales baratos. Pero, primero, si los pobres recibieran beneficios en efectivo antes que beneficios en especie, que valen menos para los pobres que los equivalentes en efectivo, tendrían más dinero para contratar abogados del sector privado cuando los necesitaran. Segundo, el sector privado de los abogados podría ser más apto para descartar las reclamaciones malas que los abogados pagados por el gobierno, no sólo porque los incentivos de principal-agente están mejor alineados en el sector privado sino también porque es menos probable que los litigantes potenciales traten de promover juicios legales dudosos si deben pagar un abogado con sus propios recursos.

XVII.8. Redistribución de la riqueza por reglas de responsabilidad: el caso del cumplimiento del código de vivienda

La asistencia de la vivienda para los pobres es un servicio tradicional proporcionado por el gobierno. Se emplean diversos métodos de asistencia. Uno de ellos consiste en contratar contratistas para que construyan casas de departamentos que el gobierno habrá de poseer y operar, rentando departamentos a los pobres a rentas nulas o nominales ("vivienda pública"). Otro consiste en dar dinero a los pobres, quienes pueden gastarlo sólo en la vivienda. Este método de subsidio tiene la característica, atractiva para un economista, de que preserva un mercado privado de la vivienda; sólo que otorga a los pobres más demanda efectiva para la vivienda proporcionada en ese mercado. Es cierto que el efecto a corto plazo puede ser simplemente el de elevar las rentas, y así empobrecer a otros rentistas, y enriquecer a los arrendadores, a resultas del subsidio. Ésa es la respuesta tradicional a corto plazo ante un aumento de la demanda, reflejando el hecho de que la capacidad productiva está por definición fija a corto plazo (véase una explicación más amplia más adelante, en XXII.15). Pero a largo plazo aumentará el acervo de vivienda y bajará el precio de la renta hacia su nivel anterior al subsidio (si bien no hasta llegar al extremo, a condición de que la curva de oferta a largo plazo observe una tendencia ascendente).

[36] En el caso de un abogado contratado para tramitar un divorcio, este efecto podría verse contrarrestado por el hecho de que el cónyuge también tiene derecho al servicio legal gratuito.

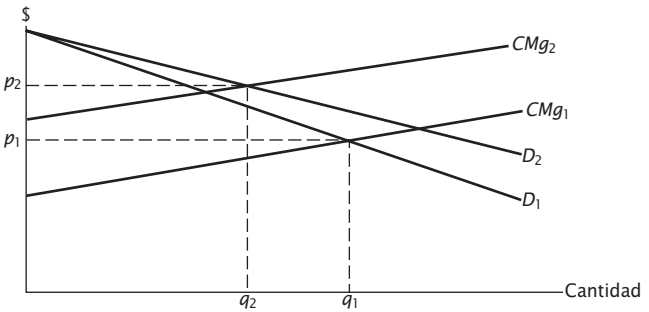
Y la vivienda pública tiene efectos de corto plazo aún peores que éstos: a corto plazo —antes de que la vivienda se construya y rente efectivamente—, los pobres no obtienen ningún beneficio de la vivienda pública. Un sistema de complementos de rentas o de vales de vivienda no sólo es más flexible, sino que opera con mayor rapidez para ayudar a los beneficiarios deseados.[37]

Tanto la vivienda pública como los complementos de la renta involucran a las ramas de ingresos y egresos del gobierno y no a los tribunales. Pero hay un método de ayuda (presunta) a los pobres para que satisfagan sus necesidades de vivienda que involucra centralmente a los tribunales: la aplicación de los códigos de vivienda. Estos códigos especifican estándares mínimos de la calidad de la vivienda, aunque se debate si ello se hace para asegurar un nivel mínimamente decente de seguridad y salubridad o para subsidiar a la industria de la construcción. En la gráfica XVII.3 se describen los efectos de la aplicación forzosa del código de vivienda. D_1 es la curva de demanda del mercado para la vivienda de los pobres antes de la aplicación. Tiene pendiente negativa porque no todos los inquilinos se marcharían si las rentas aumentaran a resultas de un aumento de los costos marginales de los arrendadores. CMg_1 es la curva del costo marginal de los arrendadores antes de la aplicación y tiene pendiente positiva para reflejar el hecho de que la creación de viviendas de rentas bajas implica el uso de ciertos recursos especializados —la tierra en particular— que valdrían menos en cualquier otro uso.

La aplicación forzosa del Código de la Vivienda tiene dos efectos principales.[38] Al mejorar la calidad de las unidades de vivienda, aumenta su demanda. Y al aumentar los costos de mantenimiento de los arrendadores, que son costos marginales porque varían con el número de las unidades de vivienda proveídas, se eleva la curva del costo marginal. El viraje mostrado en la gráfica XVII.3 es grande en relación con el desplazamiento de la curva de demanda, según el supuesto plausible de que si la cantidad demandada fuese muy sensible a un aumento de la calidad de la vivienda proveída, los arrendadores mejorarían voluntariamente la calidad y no habría necesidad de aplicar forzosamente un código de vivienda. En la gráfica XVII.3 se describen la demanda y la oferta como muy elásticas, según los supuestos (de nuevo plausibles) de que los habitantes de los barrios miserables carecen de los medios para pagar

[37] Para un análisis empírico del intercambio de pérdidas sufridas y ganancias cosechadas con la vivienda pública y los complementos de rentas, véanse Edgar O. Olsen y David M. Barton, "The Benefits and Costs of Public Housing in New York City", 20 *J. Pub. Econ.*, 299 (1983); G. Thomas Kingsley, "Housing Vouchers and America's Changing Housing Problems", en *Privatization and Its Alternatives,* 115 (William T. Gormley, Jr. [comp.], 1991); Todd Sinai y Joel Waldfogel, "Do Low-Income Housing Subsidies Increase the Occupied Housing Stock?", 89 *J. Pub. Econ.,* 2137 (2005).

[38] Un tercer efecto es la disminución de la renta de la tierra recibida por los terratenientes (suponiendo que son los propietarios de la tierra). Aquí la ironía es que estos "rentistas" incluyen cierto número de personas casi pobres para quienes la propiedad de inmuebles miserables representa la primera etapa del escape de la pobreza.



GRÁFICA XVII.3

rentas más elevadas y de que las rentas están ya tan deprimidas en relación con los costos que una nueva reducción de ellas haría que muchos arrendadores se retiraran del mercado de vivienda barata (por ejemplo, entregando sus inmuebles al municipio).

Dados estos supuestos, la aplicación forzosa del código de vivienda conduce a una disminución sustancial de la oferta de vivienda barata (de q_1 a q_2), aunada a un aumento sustancial del precio de la oferta restante (de p_1 a p_2). El efecto de la cantidad está efectivamente subestimado (aunque por la misma razón el efecto del precio está sobrestimado) en la gráfica XVII.3: parte de la oferta de mayor calidad provocada por la aplicación forzosa del código de vivienda podría rentarse a quienes no son pobres.[39] Estos efectos podrían ser contrarrestados por los complementos para la renta (pero ello privaría al programa de su calidad políticamente atractiva de no implicar gastos públicos). Si la oferta de vivienda fuera muy inelástica, el principal efecto de los complementos para la renta sería incrementar las rentas económicas recibidas por los propietarios de las casas en vez de expandir la oferta. No obstante, aparejar la aplicación del Código de la Vivienda con complementos a la renta mejoraría la calidad de la vivienda sin reducir la oferta, suponiendo que los complementos compensaran plenamente a los propietarios de las casas de los costos en que incurrieron al cumplir con el código.

Es cierto que la magnitud de los efectos mostrados en la gráfica XVII.3 depende de la ubicación (arbitraria) de las curvas. Se ha sugerido incluso que la demanda podría ser perfectamente elástica en la región relevante (lo que implica que la aplicación forzosa del código de vivienda no tiene ningún efecto sobre el precio, si bien todavía habría un efecto sobre la cantidad), por-

[39] ¿Podría ser un propósito oculto de los códigos de vivienda el de aumentar la oferta de vivienda de ingreso medio a costa de los pobres? *Cf.* George J. Stigler, "Director's Law of Public Income Redistribution", 13 *J. Law & Econ.*, 1 (1970). Sí. La "Ley de Director" (denominada así por George Stigler en honor al economista Aaron Director, quien descubrió la "ley") establece que la clase media, debido a que ejerce una gran influencia política, tiene la habilidad de redistribuir la riqueza procedente de ricos y pobres hacia sí misma.

que el aumento más ligero haría que muchos inquilinos se amontonaran.[40] Pero, dado que el amontonamiento es costoso (implica renunciar al valor del mayor espacio y la privacidad de la ocupación de una sola familia), los inquilinos estarían seguramente dispuestos a pagar *algo* para evitar verse obligados a amontonarse, siendo ese algo una renta un poco mayor. Esto implica una demanda menos que perfectamente elástica. Los datos empíricos sugieren que la gráfica XVII.3 provee una aproximación a las condiciones reales del mercado de la vivienda miserable mejor que un modelo que suponga la elasticidad perfecta de la demanda.[41]

Otra sugerencia es que la aplicación forzosa de un código de vivienda aumentaría la oferta de vivienda para los pobres si los responsables del cumplimiento concentraran sus esfuerzos en los edificios que los arrendadores estuvieran "ordeñando".[42] La ordeña se refiere a la práctica de mantener un edificio con estándares más bajos de los que serían apropiados si el arrendador intentara mantener el edificio en operación indefinidamente. Por ejemplo, el arrendador podría haberse imaginado que, debido a un cambio del barrio o a la elevación de los costos de los combustibles, probablemente tendría que abandonar (no vender) el edificio en cinco años aunque no quisiera: para ese momento sus costos variables excederían a sus ingresos por concepto de rentas. De ser así, podría tener sentido para él disminuir sus gastos en mantenimiento de inmediato, ya que todos los efectos de tales gastos a largo plazo en la preservación del edificio tendrían escaso o ningún valor para él dado el supuesto de que al final de cinco años él no podrá venderlo y tendrá que abandonarlo. Disminuyendo tales gastos, el arrendador reduciría la calidad de la vivienda, y su ingreso de rentas bajará, pero muy posiblemente menos de lo que disminuyan sus gastos en mantenimiento. Un efecto incidental podría ser que abandone el edificio más pronto aún (aunque esto pondrá fin a su corriente de ingresos más pronto), ya que esos gastos habrían mantenido al edificio en operación por un poco más de tiempo. Si la aplicación forzosa del código local de vivienda impide que el arrendador economice

[40] Bruce Ackerman, "Regulating Slum Housing Markets on Behalf of the Poor: Of Housing Codes, Housing Subsidies and Income Redistribution Policy", 80 *Yale L. J.,* 1093 (1971), criticado en Neil K. Komesar, "Return to Slumville: A Critique of the Ackerman Analysis of Housing Code Enforcement and the Poor", 82 *Yale L. J.,* 1175 (1973). Un tribunal ha sugerido que podría prohibirse a los arrendadores, según una teoría del desalojo vengativo, que abandonaran los edificios como una alternativa al cumplimiento del código si fueran "capaces" de cumplir. Robinson *vs.* Diamond Housing Corp., 463 F.2d, 853, 869 (D.C. Cir., 1972). Tal prohibición haría disminuir la oferta de vivienda para los pobres a largo plazo al incrementar los costos a largo plazo para los arrendadores o terratenientes, dado que se verían privados de la opción de abandonar la propiedad que dejara de poseer valor.

[41] Werner Z. Hirsch, *Law and Economics: An Introductory Analysis,* 71-81 (2ª ed., 1988); Hirsch, "Effects of Habitability and Anti-Speedy Eviction Laws on Black and Aged Indigent Tenant Groups: An Economic Analysis", 3 *Int'l Rev. Law & Econ*., 121 (1983).

[42] Duncan Kennedy, "The Effect of Warranty of Habitability on Low Income Housing: 'Milking' and Class Violence", 15 *Fla. State U. L. Rev.,* 485 (1987).

de este modo, es posible que demore el abandono, ya que los gastos que se ve obligado a hacer bien podrían tener algún efecto sobre la preservación del edificio, así sea inadvertidamente. Con todo, ya que los costos del cumplimiento del código son costos variables, es más probable que causen que una disminución en el ingreso de la renta y el aumento de los costos variables se intersequen más pronto, conduciendo a un abandono más temprano. Es dudoso que un tribunal o una legislatura o una dependencia de la vivienda pudieran identificar los edificios donde la aplicación forzosa del código de vivienda demoraría el abandono por su efecto contrario a la "ordeña", en lugar de acelerarlo por su efecto de encarecimiento de la continuación de la propiedad del edificio.

Varias ciudades han promulgado ordenanzas destinadas a proteger a los inquilinos otorgándoles más derechos procedimentales en el caso de que el arrendador trate de echarlos, permitiendo que retengan la renta si los arrendadores no hacen las reparaciones requeridas por el contrato, exigiendo que los arrendadores paguen intereses sobre los depósitos de seguridad, etc.[43] Los efectos son muy semejantes a los de la aplicación obligatoria del código de vivienda: las ordenanzas aumentan los costos de los arrendadores y, por lo tanto, aumentan los niveles de las rentas y disminuyen la oferta de vivienda (especialmente la vivienda de renta, porque tales leyes estimulan la conversión a cooperativas y condominios). Desde el punto de vista de la protección de los pobres, las disposiciones referentes a los derechos procesales y a la retención de la renta son particularmente perniciosas. Son derechos que invocarán más probablemente los pobres que los ricos; por lo tanto, otorgan a los arrendadores un incentivo adicional para buscar inquilinos más ricos, quienes tienen menores probabilidades de demorarse en el pago de la renta o de abusar del derecho de retenerla.

El lector podría recordar nuestro análisis, en el capítulo IV, de los efectos económicos de declarar ilegales los métodos eficientes, pero a veces opresivos, del pago forzoso de las deudas. Ambos análisis sugieren que es probable que el uso de las reglas de la responsabilidad o de otras sanciones legales para redistribuir el ingreso de los ricos a los pobres conduzca a un fracaso. Una regla de responsabilidad es como un impuesto específico: induce una contracción de la producción y un aumento del precio (véase después XVIII.3). La parte considerada responsable, aunque no sea pobre (lo que sin embargo es posible, como hemos visto), podría ser capaz de trasladar gran parte del costo de la responsabilidad a los pobres a través de precios más altos. El resultado podría ser una redistribución caprichosa del ingreso y la riqueza dentro de la clase de los pobres, y una disminución general de su bienestar.

[43] Véase Chicago Board of Realtors, Inc., *vs.* City of Chicago, 819 F.2d, 732, 741-742 (7th Cir., 1987).

XVII.9. Beneficios abiertos en especie

La ayuda a los pobres con vivienda y nutrición, y aun educación, implica costos sustanciales pero bastante previsibles, porque todos éstos son bienes bastante convencionales. Sin embargo, en el caso de la salud, sólo el cielo fija un límite potencial a los gastos. Consideremos el requerimiento de la ley federal (la Ley de Educación para Individuos con Incapacidades) de que todo niño discapacitado reciba una "educación pública apropiada y gratuita". Si el niño tiene graves deficiencias físicas o mentales, el costo podría ser astronómico (y, por esta razón, el programa no se limita a los necesitados). Podría afirmarse que cierto grado de asistencia para las necesidades educativas de los discapacitados es un método eficiente de la seguridad social contra una desgracia estrujante, y un método que implica problemas de azar o riesgo moral mínimos,[44] pero es obvio que el gasto óptimo en la educación de los niños discapacitados no es infinito ni mucho menos, aunque se requerirían gastos infinitos para llevar a esos niños hasta el nivel del niño no discapacitado. Pero ¿cuál sería el tope? ¿Cómo pueden ser racionados los servicios médicos más allá del precio? El enfoque eficiente de la cuestión de cómo decidir cuánto gastar en la educación de un niño discapacitado sería gastar hasta el nivel en el que el último dólar invertido incrementara la utilidad (bienestar, felicidad) del niño por un dólar, o, en forma alterna, gastar hasta el nivel en el que el último dólar aumente la producción económica esperada del niño por un dólar.

XVII.10. La disyuntiva de los beneficios por derecho

En 2012, el gobierno federal gastó un total de alrededor de 1.5 billones de dólares en los programas de seguridad social y de Medicare y Medicaid, los cuales financian las pensiones y la asistencia a las personas de la tercera edad. Eso ascendió a aproximadamente 40% del total del presupuesto federal. El gasto en estos programas es un gasto "por derecho", lo que significa que está financiado de manera automática en lugar de tener que ser nuevamente autorizado por el Congreso cada año, como lo son los gastos de defensa (y, de hecho, la mayoría de los gastos gubernamentales que no son "por derecho"), y lo que asimismo significa que sus receptores juzgan tener derecho a esos beneficios del programa en vez de simplemente tener la expectativa. El costo de los componentes de la pensión y la atención a la salud se está incrementando

[44] Es improbable que la perspectiva de la asistencia pública induzca a muchos padres a hacer menos esfuerzos por evitar la procreación de hijos con discapacidad. Aun la asistencia más generosa no va más allá de contrarrestar una pequeña fracción de la desutilidad que representa para el hijo y para sus padres una minusvalía severa. Esa discapacidad dificulta incluso el proceso conducente a convertir el dinero, la educación, etc., en una utilidad.

en términos absolutos y como porcentaje del presupuesto destinado al creciente porcentaje de la población que corresponde a ancianos. El gasto en atención a la salud está aumentando cada vez con más velocidad que el gasto en pensiones debido a los aumentos en el costo derivados de la introducción de nuevas tecnologías de tratamiento médico y, asimismo, debido a la curva de oferta de inclinación ascendente que resulta normal, lo que provoca que un incremento en la demanda de atención a la salud de la tercera edad ocasione un incremento en los costos promedio y, por ende, totales de la atención a la salud.

Los programas por derecho son financiados mediante impuestos a la nómina, pero la cobertura de los gastos está excediendo con rapidez a los ingresos fiscales. El gobierno tendrá que rellenar el hueco recortando los beneficios, aumentando los impuestos o recurriendo a un préstamo continuo y oneroso. De ahí que a los costos por derecho pueda tener que agregarse un costo cada vez mayor del servicio de la deuda gubernamental. Los costos correspondientes al servicio de la deuda pueden crecer más raudamente que la cantidad prestada en virtud de que, entre más pide prestado el gobierno, más elevada es la tasa de interés que probablemente tendrá que pagar. En lugar de pagar 5% para poder pedir prestados mil millones de dólares, incurriendo así en un gasto por intereses de 50 millones anuales, el gobierno podría tener que pagar un interés de 7% para pedir prestados dos mil millones, sumando de este modo un gasto al año por intereses de 140 millones: más del doble del gasto por intereses sobre un préstamo que es sólo del doble.

En 1983 el Congreso hizo una enmienda a la ley de seguridad social para estipular que, para las personas nacidas en 1938, la edad de elegibilidad para los beneficios plenos de seguridad social aumentaría gradualmente de 65 a 67 años. Así pues, los primeros efectos de la reforma no se hicieron sentir sino hasta 2003, cuando las personas nacidas en 1938 alcanzaron los 65, y el efecto pleno no se sentirá sino hasta 2026, cuando las personas nacidas en 1959 alcancen los 67; fue el acto de diferir el golpe lo que hizo al programa políticamente factible. Hoy en día no existe un movimiento en pos de una reforma similar, la cual pudiera en el futuro, por ejemplo, elevar la edad de derecho a beneficios plenos de seguridad social a los 70 años, en gesto de reconocimiento de los incrementos continuos en longevidad, salud e ingresos. En esta cuestión, el reverso de la moneda es que muchas de las personas más necesitadas de seguridad social ya están a fines de sus 60 años y encaran problemas de salud importantes; si se las hace inelegibles para beneficios de seguridad social para retiro, es probable que busquen beneficios de seguridad social por discapacidad, que son equivalentes. Con todo, el procesamiento de solicitudes por discapacidad es mucho más costoso que el de solicitudes de beneficios de seguridad social para retiro, pues este último depende de poco más que un comprobante de la edad y la historia laboral.

Hasta ahora, la nación no quiere ver la realidad en lo que respecta a sus problemas fiscales, aun cuando éstos son más severos de lo que eran en 1983.

Lecturas recomendadas

Anderson, Martin, *Welfare: The Political Economy of Welfare Reform in the United States* (1978).
Friedman, Milton, *Price Theory*, cap. 12 (1976).
Handbook of Central Banking, Financial Regulation and Supervision After the Financial Crisis (Sylvester Eijffinger y Donato Masciandaro [comps.], 2011).
Hochman, Harold M., "Contractarian Theories of Income Redistribution", en *Social Policy Evaluation: An Economic Perspective,* 211 (Elhanan Helpman, Assaf Razin y Efraim Sadka [comps.], 1983).
———, y James D. Rodgers, "Pareto Optimal Redistribution", 59 *Am. Econ. Rev.,* 542 (1969).
Osberg, Lars, *Economic Inequality in the United States* (1984).
Page, Benjamin I., *Who Gets What from Government* (1983).
Sen, Amartya K., *Collective Choice and Social Welfare*, cap. 9 (1970).
Simons, Henry C., *Personal Income Taxation*, cap. 1 (1938).
Stigler, George J., "Director's Law of Public Income Redistribution", 13 *J. Law & Econ*., 1 (1970).
Symposium, "The Top 1 Percent", 27 *J. Econ. Perspectives*, 3, verano de 2013.
Winter, Ralph K., Jr., "Poverty, Economic Equality, and the Equal Protection Clause", 1972 *S. Ct. Rev.,* 41.

Problemas

1. Si se acepta el principio de la utilidad marginal decreciente del dinero, ¿se sigue de allí que una corporación que obtenga un monopolio y así aumente sus ganancias se preocupará menos por la minimización de sus costos? ¿Hasta qué punto es aplicable el principio a las organizaciones?

2. Es usted un asistente de un senador de los Estados Unidos. El senador le pide que le aconseje si debe votar a favor de ciertos proyectos de ley que están pendientes. Desea saber sus opiniones sobre los méritos, no la política, de los proyectos. Aconséjelo. Los proyectos son los siguientes:

a) Un proyecto para mejorar la seguridad en las minas de carbón que requiere la instalación de instrumentos de alarma temprana contra explosiones, así como inspecciones de seguridad más frecuentes en las minas. El costo anual de estas medidas se estima en 500 000 dólares y se sufragará con un impuesto especial sobre las compañías mineras.

b) Un proyecto para proveer un subsidio anual, pagado de las recaudaciones fiscales generales, al servicio de pasajeros del ferrocarril en

rutas de menos de 130 kilómetros, con un monto de 10 millones de dólares por año.

3. Evalúe el siguiente pasaje:

> La distribución de la riqueza […] es un asunto de las instituciones humanas solamente […] Las reglas por las que se determina son las que quieran las opiniones y los sentimientos de la parte gobernante de la comunidad, y son muy diferentes en épocas y países diferentes; y podrían ser más diferentes aún si la humanidad así lo decidiera. [John Stuart Mill, *Principles of Political Economy* [*Principios de economía política*], libro II, cap. 1, 1, en 200 (W. J. Ashley [comp.], 1926)].

4. ¿Es probable que un programa de vales de comida aumente la nutrición de los pobres? Considere el siguiente ejemplo hipotético: una familia pobre, antes del programa de vales de comida, gastaba 700 dólares al año en alimentos. La familia recibe vales de comida con un valor al menudeo de 500 dólares. ¿Esperaría que esto afectara los hábitos alimenticios de la familia en mayor medida que una transferencia en efectivo de 500 dólares? ¿Podría verse afectada su respuesta por las restricciones que se impusieran al uso de los vales, es decir, por los alimentos específicos que no pudieran comprarse con ellos?)

5. Contraste desde un punto de vista económico los siguientes métodos para aumentar el empleo de los pobres: un programa de capacitación para el empleo subsidiado por el gobierno; un subsidio para los empleadores de trabajadores pobres; un impuesto negativo al ingreso con una tasa marginal baja.

6. ¿Es probable que el control de las rentas, como el que existe en la ciudad de Nueva York, ayude o perjudique a los pobres en conjunto? ¿Cuál sería el efecto sobre la oferta de viviendas si el control de las rentas se combinara con una aplicación estricta de las disposiciones del código de vivienda? *Cf.* Edgar O. Olsen, "An Econometric Analysis of Rent Control", 80 *J. Pol. Econ.*, 1081 (1972).

7. El profesor Ackerman (véase antes XVII.8, nota 40) sostiene que es antieconómica la construcción de viviendas baratas ahora sin un subsidio gubernamental. Si así ocurre, ¿apoya esto su propia conclusión de que es improbable que un cumplimiento estricto de los códigos de vivienda afecte la oferta de viviendas pobres, excepto por causar el abandono de algunas propiedades marginales?

8. Javins *vs.* First National Realty Corp., 428 F.2d, 1071 (D.C. Cir., 1970) sostuvo que hay una garantía de habitabilidad, medida por los estándares establecidos en el código de vivienda del distrito de Columbia, que está implicada por la operación de la ley en todos los arrendamientos. El juicio era la demanda de posesión entablada a un arrendador que alega la falta de pago

de la renta. El tribunal sostuvo que el jurado debería determinar cuál parte de la obligación del inquilino de pagar la renta, si acaso, se suspendía por el incumplimiento del arrendador. En el curso de su opinión, el tribunal sostuvo (*ibid.,* en 1079, sin las notas de pie de página) lo siguiente:

> La desigualdad del poder de negociación entre el arrendador y el arrendatario ha sido bien documentada. Los arrendatarios tienen muy poco poder para hacer cumplir las demandas de una vivienda mejor. Diversos impedimentos de la competencia en el mercado de viviendas de renta, tales como la discriminación racial y clasista y los contratos de arrendamiento de forma estandarizada, significan que los arrendadores colocan a los arrendatarios en una situación de tómalo o déjalo. La escasez cada vez mayor de vivienda adecuada aumenta más aún el poder de negociación del arrendador y agrava la necesidad de mantener y mejorar el acervo existente.

¿Qué quiere decir el tribunal cuando habla de escasez y de poder de negociación? Suponiendo que la declaración citada es correcta y coherente, ¿esperaría que la decisión del tribunal mejorara la situación de los inquilinos en general, o que la empeorara, en contraposición a que no se reconociera ninguna garantía de habitabilidad implicada? ¿Qué diremos de los inquilinos particulares que promovieron el juicio? ¿Podrían ser las rentas más elevadas, pero esas rentas más altas podrían ser compensadas por costos menores que debieran cubrir los inquilinos por tener que reparar las partes desgastadas de su lugar de alojamiento? Véase Michael A. Brower, "Comment. The 'Backlash' of the Implied Warranty of Habitability: Theory *vs.* Analysis", 69 *DePaul L. Rev.,* 849 (2011).

9. Discuta la proposición siguiente: Las políticas diseñadas para maximizar la tasa de crecimiento económico de la sociedad tienen las mayores probabilidades de satisfacer el criterio de Rawls para la justicia distributiva.

10. Supongamos que se creyera que la teoría general de la justicia de Rawls se satisface suponiendo que la gente en la posición original desearía maximizar su utilidad esperada. ¿Maximizarían la riqueza las políticas que podemos deducir de este supuesto (suponiendo que de éste pudieran deducirse algunas políticas específicas)? Véase Richard A. Posner, *The Economics of Justice,* 99-101 (1981).

11. Desde la perspectiva de un grupo de interés, ¿son las transferencias en efectivo o los pagos en especie un enfoque más atractivo de la ayuda a los pobres?

XVIII. LA TRIBUTACIÓN[1]

XVIII.1. Tributación y eficiencia

La tributación es en ocasiones un método de regulación que, como sabemos, responde al propósito de modificar la asignación de recursos (véanse antes XIV.5 y XIV.9) o la distribución de la riqueza, pero principalmente se utiliza para pagar los servicios públicos. No obstante, trae consigo repercusiones tanto en materia de asignación como de distribución, en las cuales nos concentraremos en el presente capítulo, si bien igualmente se discuten aquí los impuestos regulatorios (véase XVIII.3 después). Los impuestos sucesorios se abordan en el siguiente capítulo. De la complejidad del código fiscal nos ocuparemos con amplitud en XXI.3, más adelante; allí se hará referencia también a algunas cuestiones de recaudación. Los problemas de política fiscal que derivan de la estructura federal de nuestro sistema de gobierno se analizan en el capítulo XXVII.

Un impuesto al ingreso eficiente sería uno que obligara al usuario de un servicio público a pagar los costos de oportunidad de su uso. Pero esto trataría a los servicios públicos sólo como bienes privados, mientras que son servicios públicos precisamente porque su venta se juzga inviable o inconveniente. En el caso de algunos servicios públicos, los problemas del gorroneo impiden el uso del mercado para proveer la cantidad óptima del servicio: el individuo que se negara a pagar su parte del costo de la defensa nacional recibiría de ella la misma protección que quienes la pagaron. En el caso de otros servicios públicos, tales como la educación, la provisión del servicio por el gobierno se basa primordialmente en una determinación de que la falta de disposición para pagar el costo del servicio no debiera privar de su uso al individuo.[2]

El hincapié que se hace en las consideraciones distributivas en las discusiones de la política tributaria explica por qué aparece en esta parte del libro el capítulo sobre la tributación. Pero las consideraciones de la eficiencia también poseen gran envergadura. El hecho de gravar una actividad crea un incentivo para sustituirla por otra actividad que esté menos gravada. Se supone, sin embargo, que los individuos que decidieron hacer un viraje estaban más productivamente empleados en la primera actividad, porque de otro modo el establecimiento de un impuesto no habría sido necesario para inducirlos a

[1] Véase Boris Bittker & Lawrence Lokken, *Federal Taxation of Income, Estates, and Gifts* (3ª ed., 1999); Marvin A. Chirelstein & Lawrence Zelenak, *Federal Income Taxation* (12ª ed., 2011).

[2] Recuérdese nuestra discusión (antes, en V.4) del potencial conflicto de intereses entre padres e hijos.

cambiarse a la segunda actividad.[3] Por lo tanto, el impuesto ha reducido la eficiencia con la que se están empleando los recursos. La ineficiencia podría haberse evitado, o por lo menos disminuido, sin pérdida de recaudaciones, si el impuesto se hubiese diseñado para minimizar los efectos de sustitución. Desafortunadamente, los esfuerzos que se hagan para minimizar la ineficiencia de la asignación podrían entrar en conflicto con las metas distributivas de la política fiscal. Por ejemplo, un impuesto federal plano per cápita para adultos tendría efectos de sustitución mínimos (aunque no nulos, porque induciría a algunas personas a emigrar), pero sería muy opresivo para los pobres, a menos que fuese muy bajo… en cuyo caso generaría escasa recaudación.

En este renglón, el punto más general es que los efectos de asignación errónea que produce la tributación son minimizados fijando un gravamen a bienes o servicios o actividades para los que la demanda es inelástica (el ejemplo de Adam Smith era un impuesto a la sal). Un impuesto de esa índole no habrá de generar muchas sustituciones, y la sustitución no sólo es ineficiente sino que también reduce la recaudación obtenida mediante el impuesto. Quienes abogan por un incremento en los impuestos invariablemente hacen caso omiso de los efectos de sustitución, con lo que exageran el monto de la recaudación que esos aumentos habrían de alcanzar. En cambio, a los defensores de las reducciones fiscales les asiste la razón al hacer hincapié en que, debido a los efectos de sustitución, una cierta reducción en el porcentaje de las tasas fiscales tendrá por resultado una reducción más pequeña del porcentaje de recaudaciones fiscales, puesto que, al contar con una tasa fiscal más baja, los contribuyentes optarán por dedicarse más bien a la actividad gravada. La reducción de la tasa fiscal puede incluso determinar un aumento en las recaudaciones totales generadas por el impuesto.

A lo largo de este capítulo deberemos tener en mente un concepto útil: el de la elasticidad cruzada de la demanda, o sea, el efecto que tiene sobre la demanda de un producto (o servicio) un cambio en el precio de otro producto (véase antes X.7). Una alta elasticidad cruzada de la demanda entre dos productos significa que tales productos son sustitutos cercanos, en cuyo caso un impuesto sobre uno de ellos tenderá simplemente a desplazar a los usuarios hacia el otro, distorsionando la asignación eficiente de los recursos al mismo tiempo que genera escasa recaudación. Consideremos el tratamiento fiscal apropiado del hecho de "ir en corto contra la caja". Este término se refiere a la venta en corto de la misma acción que uno posee, eliminando así su exposición a los cambios del precio, y, habiendo eliminado así el riesgo de pérdida

[3] Esta suposición puede ser cuestionada. Es posible que haya demasiado de la actividad gravada desde el punto de vista de la eficiencia, en cuyo caso gravarla podría mejorar la eficiencia, aunque el propósito del impuesto no sea regulador. De igual modo, una deducción del impuesto podría resultar apropiada cuando no haya suficiente de la actividad desde el punto de vista de la eficiencia, aunque de nuevo es posible que la corrección de la mala asignación no sea el propósito del impuesto. Después, en XVIII.9, se discute un ejemplo.

(junto con la perspectiva de ganancia), se le permitirá generalmente retirar 95% del dinero obtenido de la venta en corto sin pagar impuesto por la ganancia (la diferencia entre el precio de venta y lo que había pagado por la acción cuando la compró). Una operación en corto es por lo tanto una venta y también es una protección casi perfecta. Si la elasticidad cruzada entre un corto contra la caja y una venta simple es mayor que la elasticidad cruzada entre un corto contra la caja y algunas formas menos eficientes de la protección, debiera tratársele para propósitos fiscales como una venta y fijársele un gravamen; de otro modo, no debería procederse así (pues en ese caso, en el que una forma menos eficiente de la protección es un sustituto cercano de un corto contra la caja, el efecto principal de gravar a este último será el de inducir la utilización sustituta de una transacción no gravada, es decir, de una forma menos eficiente de la protección, con el resultado de que habrá una distorsión de la forma de seleccionar entre las transacciones pero escasa recaudación). Ocurrirá lo contrario si un corto contra la caja es un sustituto muy cercano de una venta, que es una transacción gravada, dado que en ese caso el abstenerse de gravar el corto contra la caja sencillamente conducirá a remplazar una venta con esa clase de transacción.[4] El ejemplo es importante para abogados y jueces porque, mientras el diseño de la estructura impositiva es una función legislativa, la clasificación de transacciones particulares para determinar su estatus fiscal es administrativa y adjudicativa.

En la planeación de un sistema tributario eficiente, los criterios a considerar no incluyen solamente la capacidad de producir recaudaciones, la distorsión causada en la asignación de recursos y los efectos distributivos de un impuesto. Otro factor lo constituye el costo que representa la administración del sistema. El Servicio de Impuestos Internos (IRS) cuenta con un presupuesto anual que sobrepasa los 12 000 millones de dólares, y eso viene a ser sólo la punta del iceberg. Los costos anuales que deben cubrir los contribuyentes para preparar las declaraciones de impuestos federales se han estimado en 100 000 millones de dólares,[5] y se ha echado mano de esta cifra para instar a una simplificación del Código de Recaudación Interna. No obstante, los ahorros potenciales que se conseguirían con la simplificación son difíciles de estimar. H&R Block obtiene unos ingresos totales de casi 3 000 millones al año por preparar declaraciones de impuestos para más de 25 millones de contribuyentes, en su mayoría personas de recursos más bien modestos y, presumiblemente, con declaraciones poco complicadas. El gasto medio para la preparación prestada a dichos contribuyentes es, pues, de más de 110 dólares. El número total de declaraciones de impuestos federales procesadas en 2012 fue

[4] Véase David A. Weisbach, "An Efficiency Analysis of Line Drawing in the Tax Law", 29 *J. Legal Stud.*, 71 (2000).

[5] Edward P. Lazear y James M. Poterba, "Reforming Taxes to Promote Economic Growth", *The Economists' Voice,* 2006, http:// www.bepress.com/ev/vol3/issl/art3.

de casi 240 millones. Si uno da por sentado que una vez sometido el Código de Recaudación Interna a una drástica simplificación el costo medio por la preparación de cada una de las declaraciones bajaría a 100 dólares, entonces el costo total del sistema tributario de ingresos federales sería sólo de 24 000 millones. Con todo, esa suma es una crasa subestimación. Ignora el costo del tiempo que le toma al contribuyente obtener, y enviar a los encargados de preparar la declaración, lo que se requiera de información para completar el formulario. E ignora las justificaciones económicas que hay para tramitar una serie de deducciones y créditos, como las contribuciones caritativas, los gastos para la generación de ingresos y, asimismo, los gravámenes fijados en el extranjero y otros de carácter duplicativo. Calcular estos ítems involucra complicaciones inevitables, entre ellas la de cómo valuar las donaciones a la beneficencia que son en especie y la de cómo determinar en cuáles casos son los gastos de un negocio *realmente* un gasto, en vez de un ingreso disfrazado. A las deducciones y créditos económicamente eficientes se han de sumar otros que, tras habérseles conferido una aureola de deidad sacrosanta, no van a desaparecer del mapa. E, incluso si no hubiera deducciones, se darían sinuosidades al hacer el cómputo de los impuestos debidos al ingreso no salarial. Por añadidura, algún tipo de ingreso que sortea la tributación, como el ingreso de renta imputado por la posesión de una vivienda, debería ser gravado para evitar distorsiones; ahora bien, llevar a cabo esta empresa impondría costos adicionales para la preparación de declaraciones fiscales.

Por consiguiente, aun cuando es una buena idea la simplificación fiscal, los ahorros a derivarse de ella serían modestos. Con todo, hay dos medidas que tenderían a generar ahorros sin pasar por el tamiz de la simplificación, a saber: primero, no permitir que las cuotas por la preparación de declaraciones se deduzcan del impuesto al ingreso, y, segundo, reducir las tasas fiscales marginales, en virtud de que, entre más altas sean éstas, mayor será la compensación por los esfuerzos destinados a encontrar fisuras tributarias y, por ende, mayores serán los costos en que se incurra debido a tales empeños.

Algunos economistas se quejan de que enfatizar la necesidad de hacer más eficiente al sistema tributario constituye una perversión. Según señalan, entre más eficiente sea ese sistema, más grandes serán los desembolsos netos del gobierno —la diferencia entre los rendimientos del gobierno y el costo de generar esos rendimientos— en cualquier nivel del gasto. A medida que se alargue el tamaño del pastel en espera de ser repartido, más grande será la demanda planteada por los grupos de interés al gobierno de que muestre largueza, y si los programas procurados por esos grupos disminuyen en vez de aumentar el bienestar económico, esa disminución será más severa entre más numerosos y amplios sean dichos proyectos.[6] Aun así, será necesario

[6] Véase Gary S. Becker y Casey B. Mulligan, "Deadweight Costs and the Size of Government", 46 *J. Law & Econ.,* 293 (2003).

sopesar las pérdidas y ganancias habidas en el trueque entre semejante disminución y los ahorros en costos sociales derivados de tener un sistema tributario más eficiente. Y, por lo demás, no todos los programas gubernamentales son el producto ineficiente de las presiones ejercidas por los grupos de interés. Un sistema tributario más eficiente abre paso a un incremento de los ingresos del gobierno, que podrían utilizarse para la aplicación rigurosa de la ley, la defensa nacional, la protección ambiental, la educación, la investigación científica básica y otras actividades, algunas de las cuales bien pueden estar subfinanciadas desde la perspectiva del bienestar general.

XVIII.2. La conscripción

Muchos impuestos son implícitos.[7] El reclutamiento —conscripción— es uno de ellos, y de vez en cuando uno escucha conversaciones en torno a la conveniencia de devolverle su lugar de alguna forma al reclutamiento para que el sector militar sea más representativo de la población estadunidense en su conjunto. Las objeciones económicas al reclutamiento militar son dos. La primera de ellas es que otorga al gobierno un incentivo para sustituir con cantidades excesivas de fuerza de trabajo otros insumos de la defensa porque el precio de la mano de obra militar para el gobierno es menor a los costos de oportunidad del tiempo de los reclutados. Esta objeción es decisiva (por lo menos desde un punto de vista económico) en tiempos de paz, cuando la demanda del personal militar y los peligros afrontados por éste son relativamente pequeños, limitando así el monto del gasto gubernamental necesario para obtener el personal deseado. Pero los gastos necesarios para dotar de hombres a las fuerzas armadas en una guerra general sobre una base enteramente voluntaria serían muy grandes. Se requeriría un aumento sustancial de las tasas impositivas (o de la tasa de inflación, que es una forma de la tributación), lo que crearía —como veremos enseguida— efectos de sustitución muy ineficientes. Es improbable que un reclutamiento en tiempos de guerra que no se prevé (porque no se previó la guerra) tenga efectos de sustitución similares.[8] Es algo similar a un impuesto de capitación.[9]

Es incierto el balance entre el incentivo que el reclutamiento otorga a los militares para emplear demasiada mano de obra y los efectos de sustitución ineficientes de las tasas impositivas muy elevadas. Pero es probable que la primera ineficiencia sea relativamente poco importante en las condiciones

[7] Véase *The Military Draft: Selected Readings on Conscription* (Martin Anderson [comp.], 1982), especialmente la parte IV y el capítulo 41.

[8] Un reclutamiento previsto llevaría a los individuos a alterar sus elecciones educativas y ocupacionales, dependiendo del patrón de las exenciones del reclutamiento.

[9] Todavía podría evadirse —mediante la emigración, la automutilación, etc.—, pero éstas son medidas muy costosas para evitarlo, de modo que se recurre a ellas con muy poca frecuencia.

de una guerra total cuando los militares no están sujetos a una restricción presupuestaria estricta y podría ser óptima la asignación, a las fuerzas armadas, de la virtual totalidad de la población masculina joven que esté físicamente apta.[10]

La segunda objeción económica contra la conscripción es que produce una combinación subóptima de los reclutas porque no toma en cuenta las diferencias existentes entre los individuos en cuanto a los costos de oportunidad del servicio militar. (Por lo tanto, el reclutamiento saca de la fuerza de trabajo civil a demasiadas personas, y a las menos indicadas.) Es posible que A no sea un mejor soldado que B sólo porque A ha renunciado a una oportunidad de trabajo que vale 40 000 dólares anuales, mientras que B ha renunciado a una que sólo vale 12 000 dólares, pero mientras A no sea inferior los militares se sentirán indiferentes acerca del costo social mayor de su reclutamiento, toda vez que el gobierno aún estaría consiguiendo a su mano de obra sin costo (para él). Al menos en este ámbito, el gobierno no es un maximizador del bienestar social (o un minimizador del costo social), y, por lo tanto, saca ventaja de la brecha existente entre los costos privados y los sociales tal como lo haría cualquier empleador privado. Entre mayor sea la fracción de la población de jóvenes físicamente capacitados cuyo empleo resulta óptimo en las fuerzas armadas, menor será el problema de la selección de los mejores. Pero aun en la segunda Guerra Mundial se exentaba a jóvenes trabajadores de ocupaciones esenciales, y éste era un método burdo para ajustar por los costos sociales de oportunidad de los reclutas potenciales.

En la Guerra Civil, el Norte recurrió al reclutamiento, pero permitía que los conscriptos compraran su exención del servicio militar contratando a un sustituto para que sirviera en su lugar o pagando un honorario fijo. Esto no resolvía la primera objeción contra el reclutamiento —puesto que el gobierno todavía estaba haciéndose de la mano de obra que le faltaba sin tener que sufragar el costo total—, pero sí resolvía la segunda, dado que los reclutas potenciales para quienes el servicio militar representaba altos costos de oportunidad podían optar por salirse de él. El gobierno disponía de poderes tributarios bastante más efectivos durante la segunda Guerra Mundial que los que había tenido en tiempos de la Guerra Civil, de modo que había menos probabilidades de que la venta de excepciones resultara atractiva como una medida para recabar ingresos. Nótese la analogía del reclutamiento con el dominio eminente (véase antes III.6), pero también la diferencia: el dominio eminente confisca en el plano tributario los valores subjetivos de la propiedad, en tanto que el reclutamiento hace lo mismo con los costos de oportunidad del recluta.

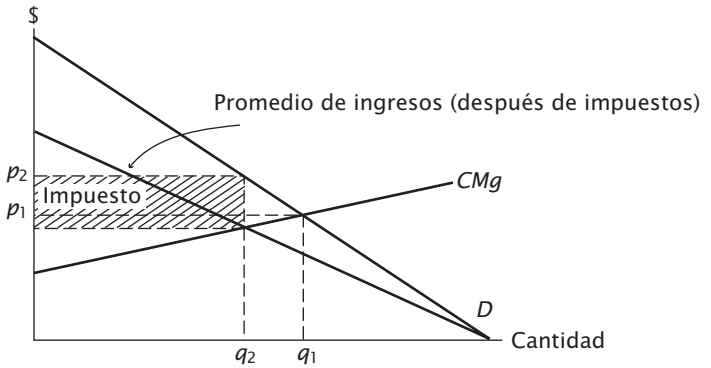
[10] A pesar del tiempo presente de esa oración, en el momento de escribir estas líneas resulta difícil concebir una guerra (por larga que sea) que pudiera requerir algún día la movilización de un ejército masivo de conscriptos a la escala de nuestras fuerzas de la segunda Guerra Mundial o que acabara con la exención de las mujeres respecto de la conscripción.

XVIII.3. IMPUESTOS ESPECÍFICOS

Empezamos nuestra consideración de los impuestos explícitos con un impuesto sobre las ventas al menudeo: por cada adminículo vendido, el productor deberá remitir al gobierno 30% del precio.[11] Se supone que la industria de adminículos es competitiva, lo que quiere decir que el precio de un adminículo era igual a su costo marginal antes de que se estableciera el impuesto. El efecto del impuesto es un desplazamiento de la curva de demanda de la industria, vista aquí como la curva de ingreso medio para la industria antes que la del costo medio para el consumidor, de pendiente negativa como se muestra en la gráfica XVIII.1. El precio de la industria se determinará por la intersección del costo marginal con el nuevo ingreso medio. Ese precio, p_2, es mayor que p_1, el precio vigente antes del establecimiento del impuesto. Al precio mayor, la cantidad demandada por los consumidores, q_2, es menor a la cantidad demandada antes del establecimiento del impuesto. El precio mayor induce a algunos consumidores a desplazarse a otros productos que ahora parecen más baratos, aunque en efecto su producción es más costosa que la de los adminículos o son inferiores en calidad a los adminículos, porque de otro modo la sustitución habría ocurrido antes de que el impuesto elevara el precio. El impuesto tiene el mismo efecto de sustitución que la fijación monopólica del precio de los adminículos. Éste es un costo de eficiencia que conlleva el impuesto.

La magnitud del efecto depende, como se sugirió antes, de la elasticidad-precio de la demanda. Si la demanda es muy elástica, un aumento relativamente pequeño del precio, inducido por el impuesto, conducirá a una reducción relativamente grande de la producción. La recaudación generada por el impuesto, siendo una función de las ventas, será también menor que si el impuesto se hubiera establecido sobre un producto cuya demanda fuese menos elástica. Por lo tanto, las metas de la minimización de los efectos de sustitución de la tributación y de la maximización de los ingresos fiscales del

[11] Para simplificar, supongamos que el productor vende directamente al consumidor. ¿Cómo cambiaría el análisis si el impuesto gravara a un distribuidor detallista que le comprara al productor y le revendiera al consumidor? En general, el nivel al que se imponga nominalmente el gravamen no implicará diferencia alguna en la forma en que tiene en último término incidencia y provoca efectos. El esfuerzo de los minoristas por desplazar el impuesto si se les carga habrá de inducir a los consumidores a causar efectos de sustitución, a la vez que traerá consigo la misma repercusión sobre la producción realizada por los productores. En efecto, toda vez que los costos de las transacciones entre el minorista y el productor rara vez serán elevados, podemos emplear el Teorema de Coase para concluir directamente que, ya sea que la responsabilidad fiscal recaiga inicialmente en el detallista o en el productor, esto no tendrá efecto en último término en la asignación de recursos (pasando por alto efectos de ingreso que son presumiblemente menores).



GRÁFICA XVIII.1

gobierno aconsejan, como se sugirió antes, que se grave en mayor medida a los bienes o servicios inelásticos respecto al precio.[12]

¿Pero qué diremos de los efectos distributivos? La recaudación generada por el impuesto se representa en la gráfica XVIII.1 por el área sombreada. Adviértase que la mayor parte del impuesto, $q_2 \times (p_2 - p_1)$, es soportada por los consumidores de adminículos. Los impuestos al consumo tienden a ser regresivos; es decir, a tomar una fracción mayor del ingreso de quienes no son ricos que del ingreso de los ricos, porque quienes no son ricos consumen una fracción mayor de su ingreso. Es cierto que la naturaleza del producto —que sea comprado más probablemente por una persona rica que por una no rica— es también importante. Es probable que un impuesto a los yates fuese progresivo (sujeto a una calificación señalada más adelante). Pero los productos fuertemente demandados por los ricos no son necesariamente (o generalmente) aquellos cuya demanda es inelástica, de modo que un impuesto específico eficiente será a menudo regresivo, y un impuesto específico progresivo será a menudo ineficiente.

La parte del impuesto al adminículo en la gráfica XVIII.1 que no se desplaza en dirección a los consumidores proviene de las rentas económicas recibidas por los propietarios de factores de producción utilizados para fabricar adminículos. Un impuesto que recayera *enteramente* sobre las rentas tendría características atractivas. Por definición, no habría ningún efecto de sustitución. Las rentas son un rendimiento por encima de los costos de oportunidad; es decir, rendimientos alternativos, de modo que una reducción de las rentas no inducirá un desplazamiento del recurso a un uso alternativo. Y un impuesto a las rentas *suena* ciertamente progresivo. Pero puede no serlo. Un efecto de

[12] Adviértase la analogía con el principio de Ramsey para la determinación de los precios de los servicios públicos (véase antes XIII.5).

nuestro hipotético impuesto a los adminículos es el de reducir la producción de la industria y, por lo tanto, presumiblemente, su demanda de insumos, incluida la mano de obra. Si los trabajadores despedidos por la industria tienen oportunidades de empleo igualmente buenas en otra parte, y costos de reubicación nulos, no se verán perjudicados por el impuesto. Pero en caso contrario perderán rentas que ganan en su empleo actual.[13]

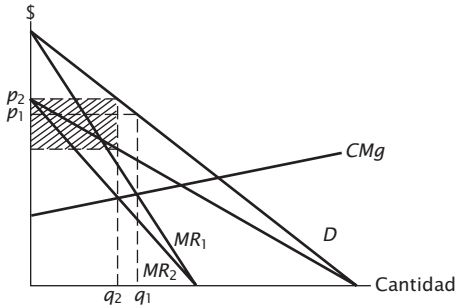
Si las ganancias monopólicas pudieran detectarse y medirse, podrían gravarse directamente y, dado que son rentas, el impuesto no tendría efectos de sustitución. (Lo mismo se aplicaría a un impuesto sobre el excedente del consumidor.) Sin embargo, un impuesto *específico* sobre los monopolios tiene esencialmente los mismos efectos de sustitución y distribución que los impuestos específicos sobre los bienes de los vendedores competitivos. Como se muestra en la gráfica XVIII.2, el impuesto específico confronta al monopolista con una nueva tabla de demanda (ingreso medio), lo que lo lleva a recalcular su lista de ingreso marginal. Su nuevo precio, en la intersección del costo marginal y el nuevo ingreso marginal, es mayor, y su producción es menor. El impuesto (el área sombreada en el diagrama) es soportado por los consumidores, el monopolista bajo la forma de menores ganancias monopólicas y los propietarios de los recursos bajo la de oferta inelástica de los insumos empleados en la fabricación del producto.

Debemos señalar otras aristas del análisis económico de los impuestos específicos que deben ser depuradas.[14] Hemos supuesto implícitamente que el impuesto es cierta fracción del precio del bien *(ad valorem)*. Pero un impuesto calculado de este modo dará a la empresa un incentivo para reducir la calidad y, por ende, el precio del bien. Por ejemplo, si el impuesto específico sobre los automóviles se basa en el precio de menudeo del automóvil con todas las opciones (aire acondicionado, estéreo, etc.), y si las opciones que se compran por separado no pagan ningún impuesto o pagan una tasa menor, el fabricante tendrá un incentivo para ofrecer menos opciones con el automóvil. En cambio, si el impuesto es plano (por ejemplo, 100 dólares por automóvil), el fabricante tendrá un incentivo para aumentar la calidad, estando exentos del impuesto los aumentos de la calidad.

Como vimos al discutir los impuestos a la contaminación en XIV.6, los impuestos específicos suelen utilizarse para propósitos de regulación. Un impuesto puramente regulatorio no generaría ningún ingreso, ya que reduciría la demanda del producto o la actividad regulados a cero. Empero, a menudo,

[13] Si no estuviéramos discutiendo un impuesto específico sobre los adminículos o los yates, o alguna otra parte menor de la actividad productiva de la nación, sino un impuesto específico general, es decir, un impuesto sobre los bienes vendidos, sería inadecuado el tipo de análisis de equilibrio parcial esbozado en el texto: un análisis donde se omiten las interacciones existentes entre la actividad gravada y el resto de la economía, en contraste con el análisis de equilibrio general.

[14] Yoram Barzel, "An Alternative Approach to the Analysis of Taxation", 84 *J. Pol. Econ.,* 1177 (1976).



GRÁFICA XVIII.2

como en el caso de los impuestos específicos a los cigarrillos o al alcohol, el propósito y el efecto consisten en reducir la demanda y generar recaudaciones a base de las ventas del producto que siguen lográndose. A veces se afirma (erróneamente) que los impuestos específicos a los cigarrillos no pueden reducir su consumo, porque el hábito de fumar es una adicción. Por el contrario, un impuesto específico a los cigarrillos debiera tener un efecto drástico sobre el consumo de los adictos al cigarrillo. Un adicto, en términos económicos, es alguien cuya demanda del producto adictivo es una función positiva de su consumo anterior. (La adicción es así una forma de la adaptación.) El adicto racional percibirá un aumento del costo del producto como un aumento de sus gastos futuros antes que simplemente de sus gastos presentes, porque el consumo mayor del producto ahora hará más probable que lo compre en el futuro (véase antes VII.11). El valor presente de esos costos futuros deviene una razón adicional para dejar el cigarro en el momento actual.

Podría creerse también, y de nuevo erróneamente, que un impuesto específico a los cigarrillos tendría un efecto menor sobre el consumo de los pobres y escasamente educados (hoy en día son en gran medida las mismas personas) que sobre el consumo de los ricos. Es menos probable que los pobres aprecien las consecuencias del hábito de fumar para la salud, y, por lo tanto, su demanda de cigarrillos será menos elástica que la de los ricos. Es cierto lo contrario. El precio completo de los cigarrillos incluye el precio de compra más cualesquiera costos asociados al uso del producto, primordialmente los riesgos para la salud del hábito de fumar. El precio completo de los cigarrillos será mayor para los consumidores educados. Los costos de la mala salud son mayores, al menos en un sentido pecuniario, entre mayor sea nuestro ingreso (hay más que perder si nos incapacita la mala salud); los mejor educados tienen menores costos de información acerca de la salud que los deficientemente educados; y el ingreso y la educación se correlacionan positivamente. Por lo tanto, la demanda de cigarrillos por parte de los consumidores educados será menos elástica al precio que la demanda de quienes tienen poca educa-

ción, porque el precio de compra es una fracción menor del precio completo de los cigarrillos para los educados y el impuesto se mide sólo por el precio de compra. De modo que el impuesto específico, que podría percibirse como un impuesto de 10% por los no educados y de sólo 5% por los educados, tendrá un impacto menor sobre las compras de estos últimos.[15]

Cuando el propósito de un impuesto específico es la corrección de una externalidad, tal como los costos de accidentes resultantes del hecho de conducir en estado de ebriedad,[16] un impuesto plano es superior a un impuesto específico *ad valorem* debido a que la externalidad es proporcional a la producción: la cantidad bebida, no el precio por botella.

La legalización, al ir aparejada con un gravamen específico, puede ser una alternativa seductora a la penalización de la mariguana, la cocaína y otras drogas que están prohibidas por alterar el funcionamiento de la mente.[17] Se ha estimado que los impuestos específicos (federales, estatales y, posiblemente, también locales) generarían un ingreso fiscal total de 46 700 millones de dólares al año.[18] No obstante, esto depende de una serie de supuestos, como el de que la demanda de drogas no se incrementaría: una suposición contraria a toda lógica, en virtud de que la ilegalidad actual de las drogas no incita a efectuar su compra. Comoquiera que sea, el incremento en el consumo de drogas derivado de la legalización podría sobrevenir en una gran proporción a expensas del consumo de alcohol, tabaco y otros artículos que son gravados tal como lo serían las drogas. Asimismo, un impuesto específico a las drogas elevaría el precio para el consumidor… aun cuando probablemente el efecto neto acarreado por la legalización sumado a la tributación consistiría en reducir el precio, ya que la legalización eliminaría los costos ingentes que la prohibición les impone a los vendedores, en particular la amenaza de ser objeto de castigo y de la violencia pandillera. Con los precios reducidos, las ventas subirían, si bien el ingreso fiscal por venta bajaría si la tasa tributaria fuera un porcentaje fijo del precio de venta. Los rendimientos fiscales adicionales se generarían merced al impuesto al ingreso pagado por los vendedores; los gánsteres de la droga no pagan un impuesto al ingreso o a partir del ingreso que obtienen por su actividad delictiva.

[15] Para confirmar las evidencias, véase Gary S. Becker, Michael Grossman y Kevin M. Murphy, "Rational Addiction and the Effect of Price on Consumption", 81 *Am. Econ. Rev. Papers & Proceedings,* 237, 240 (mayo de 1991). ¿Es el término de "adicto racional" un oxímoron? No si por adicto se define simplemente a aquel cuya demanda de determinado bien o servicio es una función positiva del consumo que hizo de él en el pasado. Uno podría ser adicto a la música clásica.

[16] Por cierto, una externalidad muy sustancial. Véase Steven D. Levitt y Jack Porter, "How Dangerous Are Drinking Drivers?", 109 *J. Pol. Econ.,* 1198 (2001).

[17] Véase Jeremy Bulow y Paul Klemperer, "The Tobacco Deal", en *Brookings Papers on Economic Activity: Microeconomics,* 323 (Martin Neil Baily, Peter C. Reiss y Clifford Whinston [comps.], 1998). La argumentación de los autores a favor de la legalización se resume antes en VII.11.

[18] Véase Jeffrey A. Miron y Katherine Waldock, The Budgetary Impact of Ending Drug Prohibition (2010), www.cato.org/pubs/wtpapers/DrugProhibitionWP.pdf.

Se espera comúnmente (expectativa no carente de interés) que un impuesto fijado a las drogas hasta hoy en día ilegales había de ser un impuesto regulatorio, esto es, un impuesto diseñado para reducir la producción. Los autores conciben a ese impuesto como un impuesto al ingreso cuyo efecto neto sería incrementar la producción del ítem gravado.

Si se suman los gastos por la aplicación de la ley que serían evitados a las recaudaciones fiscales que se generarían al combinar la legalización con el impuesto específico, se obtendría, de acuerdo con las estimaciones de los autores,[19] la cantidad impresionante de 88 000 millones de dólares anuales.

XVIII.4. Impuestos prediales

La famosa propuesta de Henry George de remplazar todos los impuestos con un solo impuesto sobre la tierra era un intento de limitar la tributación a las rentas puras.[20] Pero la renta de la tierra no es suficientemente grande para satisfacer todas las demandas de recaudación de un gobierno moderno. Hay también un problema de medición, surgido del hecho de que es igualmente probable que el propietario use la tierra o la rente a otros. Si soy propietario de un terreno y construyo en él un edificio de departamentos que rento, una parte de la renta que recibo será la renta de la tierra, pero otra parte será el ingreso derivado de las mejoras que se hagan a la propiedad, y este ingreso podría no incluir ninguna renta económica. Y, si ocupo una casa de mi propiedad en mi tierra, recibiré una renta imputada igual a la renta que podría haber obtenido rentando la tierra; pero, no siendo monetaria, es probable que la renta imputada escape al impuesto. Otro problema es que muchos terratenientes no son ricos aunque obtengan rentas de la tierra. Un impuesto pesado sobre la renta de la tierra empobrecería a los agricultores, los trabajadores, los jubilados y otras personas que hubiesen comprado tierra a un precio que capitalizara las rentas esperadas. Supongamos que un agricultor, utilizando dinero prestado, compra una buena tierra agrícola por 1 000 dólares que rinde una renta anual de 100 dólares, descontados los impuestos prediales. Más tarde se establece un impuesto de 90 dólares al año sobre la propiedad. El rendimiento del agricultor bajará a 10 dólares. Si vende la tierra, sólo recuperará una pequeña fracción del dinero que pagó por ella.

En la práctica, los impuestos prediales gravan la renta de la tierra pero mucho más que eso. En efecto, las mejoras de la tierra se gravan de ordinario más fuertemente que la tierra misma. Un resultado es que se otorga a los terratenientes un incentivo artificial para evitar (o posponer) el mejoramiento de su tierra. Más importante aún, un impuesto sobre las mejoras de la tierra

[19] Véase *ibid*.
[20] Henry George, *Progress and Poverty* (1879).

es esencialmente un impuesto específico. Consideremos una construcción de casas de departamentos. Si el impuesto predial es proporcional al monto de las rentas, el propietario lo tratará en la misma forma que nuestro fabricante de adminículos trató al impuesto a las ventas: como una reducción del ingreso medio en un porcentaje uniforme en todos los niveles de la producción. A fin de maximizar el rendimiento en estas circunstancias, el propietario aumentará el precio y reducirá la producción. La carga del impuesto será compartida entonces entre los consumidores —las personas a quienes se rentan los departamentos— y un rentista que es el propietario de la tierra.

Los impuestos prediales minan los derechos de propiedad. Supongamos que soy un agricultor en un área donde se desarrolla más y más tierra para propósitos residenciales. Mi tierra vale sólo 100 000 dólares como tierra agrícola, pero un desarrollador me ofrece 200 000 dólares. Me rehúso porque me siento sentimentalmente ligado a la tierra y no quiero mudarme; no vendería por menos de 250 000 dólares. ¿Cuál valor asignaría al inmueble el tasador fiscal? Si lo valúa al valor más alto, podría obligarme a venderlo al desarrollador, porque mi ingreso agrícola podría ser insuficiente para pagar un impuesto tasado sobre la base de un uso de la tierra que generaría un ingreso pecuniario mayor. Pero la tierra es más valiosa para mí que para el desarrollador. El impuesto predial tiene aquí el mismo efecto que el dominio eminente, que también extingue sistemáticamente los valores de la tierra en exceso del precio de mercado (véase antes III.6).

No obstante, el análisis es incompleto. El impuesto predial vuelve al gobierno en efecto un copropietario de "mi" propiedad. Esos 200 000 dólares que el desarrollador está dispuesto a pagarme son el valor de la tierra para él, sujeto al impuesto a la tasa residencial antes que a la agrícola. Mi precio de reserva de 250 000 dólares se basa en la tasa agrícola. También debe considerarse el valor de la tierra para el gobierno a las diferentes tasas impositivas, es decir, el valor presente de las dos corrientes diferentes de recaudaciones fiscales esperadas. Supongamos que, a la tasa impositiva agrícola, ese valor es de sólo 10 000 dólares, pero que a la tasa residencial —el uso más elevado y mejor desde el punto de vista del gobierno— sería de 60 000 dólares. Entonces el valor total de la propiedad aumentaría al obligarme a venderla al desarrollador. Existe una analogía con el efecto del flujo de retorno sobre la venta de derechos de agua, que exploramos en III.6: el beneficiario del flujo de retorno, como la autoridad tributaria en el caso presente, es tratado como un copropietario del inmueble.

En el ejemplo que acabamos de presentar, la existencia de una oferta de compra simplifica grandemente el problema de la valuación de inmuebles para propósitos tributarios. Evalúe la siguiente proposición para una autoevaluación: el propietario valúa su propiedad para fines de la tributación de inmuebles, pero se le obliga a aceptar cualquier oferta de compra de la propiedad al valor tasado (véase antes III.6, nota 27).

La política tributaria está fuertemente inclinada a favor de la práctica de fomentar la adquisición de una vivienda por medio de una deducción del impuesto al ingreso personal. Los impuestos a bienes raíces son deducibles, el interés hipotecario es deducible, el interés sobre un préstamo inmobiliario-de capital social es deducible; el interés sobre otro préstamo de carácter no comercial no es deducible, lo que crea un fuerte incentivo para pedir un préstamo de cara a los bienes raíces. Y los beneficios de un préstamo inmobiliario-de capital social pueden ser y usualmente son utilizados para comprar bienes y servicios no relacionados con la vivienda, con lo que se propicia el uso de una casa como garantía prendaria para lo que equivale a ser un crédito para el consumidor.

Hay una distorsión adicional en el tratamiento fiscal de las viviendas ocupadas por el propietario que se crea al permitirle al vendedor sortear la obligación de pagar el impuesto a cualesquiera ganancias de capital si reinvierte las ganancias en la compra de otra vivienda. La justificación que se da para esta provisión —evitar la tributación al incremento puramente nominal de los precios causado por la inflación— es poco convincente, según veremos al analizar la tributación de las ganancias de capital.

El tratamiento fiscal favorable de la propiedad de una casa puede parecer justificado por los beneficios externos de semejante adquisición —se argumenta que las personas cuidan mejor una vivienda de su propiedad que una que alquilan, y que esto genera externalidades positivas al crear una vecindad más agradable y más segura— y debido al costo artificialmente elevado de la construcción a resultas del carácter ineficiente de los códigos de construcción y las leyes de zonificación.[21] Pero esos costos más altos son reales y, por ello, han de ser evitados en vez de inducir a la gente a asumirlos. Y la propiedad de una vivienda genera una externalidad tanto negativa como positiva: reduce la movilidad de la mano de obra debido a que los costos de transacción por vender una casa y comprar otra elevan el costo de optar por un empleo que requerirá una mudanza. Esto resultó ser un problema significativo en la crisis económica que comenzó en el otoño de 2008 subsiguiente al estallido de la burbuja de la vivienda; las personas que perdieron su trabajo no pudieron cubrir el gasto de una mudanza, ya que no podían vender su casa sin incurrir en una cuantiosa pérdida. En efecto, la tributación y otras formas de subsidiar la propiedad de una vivienda desempeñaron un papel significativo en dicha burbuja y, por consiguiente, en la crisis desencadenada por ésta.

[21] Véase, por ejemplo, Edward A. Zelinsky, "Efficiency and Income Taxes: The Rehabilitation of Tax Incentives", 64 *Tex. L. Rev.*, 973 (1986).

XVIII.5. La tributación del ingreso corporativo

El impuesto al ingreso corporativo es en un aspecto un instrumento burdo para gravar las rentas; en este caso la renta consistente en la diferencia que existe entre los ingresos totales de una empresa y sus costos totales.[22] El problema es que el impuesto, en su forma actual, no permite una deducción del costo del capital accionario y así falla no sólo en cuanto a las ganancias en el sentido económico, u otras rentas, sino también (y quizá principalmente) en cuanto al costo del capital accionario. El resultado es un incentivo para la utilización de otras formas del capital, tales como la deuda y la inversión de la empresa en el capital humano de sus trabajadores, cuyo precio para la empresa es deducible; la sustitución de insumos de capital por insumos de mano de obra, y el empleo de otras formas de la organización empresarial en lugar de la forma corporativa. Asimismo, como vimos en el capítulo XV, el impuesto al ingreso corporativo contribuye al riesgo de la quiebra y aumenta los costos de agencia. Se ha sostenido que la resistencia a la derogación del impuesto se debe en parte a su apoyo tácito por parte de los administradores corporativos,[23] y, dado que el costo del capital accionario (por oposición a las rentas monopólicas o de otra clase) es un costo marginal, una parte del impuesto al ingreso corporativo se traslada a los consumidores bajo la forma de precios mayores del producto, justo como un impuesto específico (explícito).[24] Igualmente, el impuesto al ingreso corporativo se traduce en una tributación por partida doble: el ingreso pagado en dividendos a los accionistas es gravado primero a nivel corporativo y después a nivel individual. Aun así, se encuentra a tal punto inmerso en una red de excepciones y oportunidades para la evitación que corresponde a solamente 8% de las recaudaciones fiscales federales.

Surgen muchos problemas legales intrincados en el área del impuesto al ingreso corporativo, y con frecuencia éstos tienen una dimensión económica. Veamos un ejemplo. Si se liquida una corporación que tiene una propiedad valorada, transfiriendo así la propiedad a los accionistas, quienes la venden más tarde, ¿debiera tratarse la venta como hecha por la corporación, en cuyo caso deberá pagarse el impuesto al ingreso, o como una venta hecha por los accionistas, en cuyo caso no se debe pagar ningún impuesto al ingreso corporativo? El último fue el enfoque aceptado por el Código de Recaudación

[22] El impuesto sobre la renta corporativo sirve también para limitar la explotación, por parte de los contribuyentes, de cualquier tratamiento favorable que se otorgue a las ganancias de capital. Véase después XVIII.10.

[23] Jennifer Arlen y Deborah M. Weiss, "A Political Theory of Corporate Taxation", 105 *Yale L. J.,* 325 (1995). Véase antes XV.4.

[24] Existe una literatura empírica de carácter no concluyente acerca de la incidencia del impuesto al ingreso corporativo. Un estudio notable es el de Marian Krzyzaniak y Richard A Musgrave, *The Shifting of the Corporation Income Tax* (1963). Véase también Richard A. Musgrave y Peggy B. Musgrave, *Public Finance in Theory and Practice,* 249-269 (5ª ed., 1989).

Interna hasta 1986 y resultaba objetable por razones económicas al crear costos de transacción innecesarios (los costos de una liquidación motivada por el deseo de eludir el impuesto al ingreso corporativo). En algunos casos había resultados de mala asignación más graves aún. Supongamos que una corporación tenía una planta, que estaba plenamente asegurada, y que se incendió. El valor del seguro, menos el costo original depreciado de la planta, sería tratado como una ganancia gravable de la corporación aunque ésta decidiera liquidarse. Pero si en lugar de eso la corporación comprara o construyera una planta nueva, utilizando para este fin los beneficios completos del seguro, y luego vendiera la planta y repartiera el dinero de la venta entre sus accionistas de acuerdo con un plan de liquidación completa adoptado antes de la venta, no habría ningún impuesto al ingreso corporativo.[25] La diferencia del tratamiento fiscal podría inducir a una empresa a construir o comprar una planta nueva aunque fuese más eficiente la liquidación sin comprar o reconstruir.

Los tribunales han tratado de reducir los costos sociales de la reorganización corporativa mediante la doctrina de "la sustancia sobre la forma" elaborada por los jueces, según la cual una reorganización u otra transacción cuyo único propósito y efecto sea la elusión de impuestos será descartada para propósitos fiscales.[26] Las reorganizaciones corporativas están exentas del impuesto al ingreso a fin de permitir las transacciones que tienen consecuencias económicas potencialmente benéficas, tales como la reasignación del riesgo a quienes pueden soportarlo mejor, la reducción de los costos de agencia o el traslado de los activos a usos más valiosos. Las transacciones que no tienen esas consecuencias potenciales porque su único propósito y efecto es la reducción de los impuestos no debieran alentarse mediante un tratamiento fiscal favorable porque sólo generan costos de transacción y trasladan la carga fiscal a otros contribuyentes. Están puramente en busca de renta.

La doctrina de "la sustancia sobre la forma" ha sido criticada por su carácter esponjoso o su falta de consistencia.[27] Pero esto significa que es un estándar más que una regla. En este caso como en muchos otros, el estándar sirve de punto de apoyo o colchón a las reglas. Las reglas cobran exactitud al especificar escrupulosamente cuáles conductas están prohibidas (o permitidas), pero de esta manera se abren fisuras legales para una conducta que tiene las mismas consecuencias que la conducta precisada en la regla pero que, por no estar especificada en ésta, es por definición un asunto no cubierto en ella. El estándar atrae para sí esa conducta al evitar la especificidad, siendo esta ausencia la fuente en la que nada ingrávido es estándar. El intercambio efectuado favo-

[25] Central Tablet Mfg. Co. *vs.* United States, 417 U.S., 673, 690 (1974).

[26] Gregory *vs.* Helvering, 293 U.S., 465 (1935); Falconwood Corp. *vs.* United States, 442 F.3d, 1339, 1349 (Fed. Cir. 2005); Yosha *vs.* Commissioner, 861 F.2d, 494 (7th Cir., 1988).

[27] Joseph Isenbergh, "Musings on Form and Substance in Taxation", 49 *U. Chi. L. Rev.,* 859 (1982).

rece a la doctrina de "la sustancia sobre la forma" debido a la variedad imposible de anticipar (por los redactores de la regla) que adquieren las formas transaccionales de lograr el mismo resultado.[28]

Si se vende un negocio y sus activos tangibles se valúan a un nivel inferior al del precio de compra del negocio, la diferencia es atribuida por los contadores (con el objeto de crear un balance en los libros) al valor de la empresa como un negocio en marcha y, mientras no hubiera una enmienda al código tributario, no se podría depreciar para propósitos fiscales.[29] ¿Tiene la regla de la no depreciación sentido económico?[30] Si los activos físicos de la empresa se valúan capitalizando las ganancias que se espera que genere la empresa con ellos, entonces, dado que el precio de compra será similarmente una capitalización de las ganancias esperadas, no habrá nunca un residuo. Pero si los activos físicos se valúan sobre la base de lo que costaría su remplazo, bien podría haber un residuo. Entre las posibles razones para la discrepancia, dos parecen particularmente importantes. Primero, podría haber activos no físicos, en particular el capital humano en diversas formas. Segundo, la empresa podría tener algún poder monopólico, de modo que el precio de su producción no bajará hasta el costo marginal. En el segundo caso, otorgar una deducción fiscal dependerá, en principio, de la fuente del poder monopólico, pero sería difícil que ésa fuera una investigación viable para el Servicio de Impuestos Internos en toda venta de empresas. En el primer caso, el problema consiste en la estimación de la vida de una clase de activo poco conocido (para las autoridades fiscales) a fin de determinar el periodo de depreciación óptimo. Sin embargo, dado que la ley actual permite que las empresas depreciен sus activos físicos, para propósitos fiscales, más deprisa de lo que se espera que los activos se deprecien efectivamente, sería arbitrario preocuparse por la depreciación de los activos no físicos a una tasa más rápida que la de su depreciación efectiva. Cuando se suman a todo esto la dificultad y el costo de la valuación de activos por métodos distintos de los del mercado, podrá formularse un argumento fuerte a favor de que se permita al comprador depreciar todo el costo de adquisición del negocio del vendedor.

[28] Véase David A. Weisbach, "An Economic Analysis of Anti-Tax-Avoidance Doctrines", 4 *Am. Law & Econ. Rev.,* 88 (2002).

[29] La enmienda permite que el valor de la empresa como un negocio en marcha con buen funcionamiento se deprecie a lo largo de un periodo de 15 años. 26 U.S.C. 197. El valor de buen funcionamiento no se debe confundir con la buena voluntad, un activo intangible que crea, por ejemplo, la publicidad. El valor de buen funcionamiento no puede asignarse a un activo específico; se le puede concebir como el valor generado por la integración y la administración de los activos específicos por la empresa.

[30] Véase Richard L. Doernberg y Thomas D. Hall, "The Tax Treatment of Going-Concern Value", 52 *Geo. Wash. L. Rev.,* 353 (1984).

XVIII.6. Impuestos sobre la renta contra impuestos al consumo

Podemos definir un impuesto óptimo como aquel que: *1)* tiene una base tributaria grande, lo que facilita la recaudación de las grandes cantidades de rendimientos que demandan los gobiernos modernos; *2)* grava una actividad cuya demanda es muy inelástica, de modo que los efectos de sustitución del impuesto son mínimos; *3)* no aumenta la desigualdad[31] ni ofende a la equidad, y *4)* su administración no es cara.

El impuesto al ingreso personal parecería aproximarse a este óptimo. La base tributaria es grande, la demanda de ingreso es presumiblemente inelástica, el ingreso es una medida amplia del bienestar y la recaudación del impuesto se facilita por el hecho de que la mayoría de la gente está empleada en organizaciones. Esto aún deja de lado el considerable costo de preparar las declaraciones fiscales que se ha de considerar de cara al impuesto; y, más importante aún, las segunda y tercera aspiraciones mencionadas antes no pueden satisfacerse con un impuesto al ingreso porque dependen de una definición del ingreso tan amplia que se crean costos administrativos prohibitivos y así violan la cuarta aspiración.

Consideraremos algunas de las exclusiones del ingreso en XVIII.8, pero una exclusión, particularmente interesante para el economista, debe mencionarse de inmediato. Tal es la exclusión del ingreso no monetario generado por el ocio. El impuesto al ingreso es básicamente un impuesto al trabajo, porque el principal ingreso gravado es el de los ingresos ganados, de modo que crea un incentivo para sustituir el trabajo por el ocio, cuyo ingreso (real, aunque no pecuniario) no está gravado. Este efecto podría contrarrestarse, por lo menos en principio, gravando los bienes que son complementarios del ocio, tales como el golf. Más realistamente, el efecto podría contrarrestarse al menos en principio por lo que se designa como el "efecto de ingreso" del impuesto al ingreso. Dada una utilidad marginal decreciente del ingreso, una reducción del ingreso después de impuestos, a resultas de un impuesto oneroso al ingreso logrará, al empujar al contribuyente hacia arriba en su curva de utilidad marginal, aumentar su demanda de ingreso y así lo llevará a trabajar más de lo que sería el caso si el ingreso adicional le confiriera una utilidad menor. Los estadunidenses están acostumbrados a trabajar duro y a lograr y mantener un nivel elevado de gastos en bienes del consumidor. Por añadidura, las personas que tienen ingresos altos observan la tendencia a ser altamente competitivas y a valorarse a sí mismas en relación con los ingresos de sus pares. En virtud de que brilla por su ausencia una aristocracia y de que prevalece una actitud en general filistea hacia la cultura, en los Estados Unidos el dinero

[31] La cuestión de hasta qué punto debería el sistema tributario tratar de reducir la desigualdad económica se analiza en XVIII.11 y XIX.1 después.

es el índice clave para el prestigio y la posición social, así como, a ojos de muchos ciudadanos de este país, un factor esencial para consolidar un sentimiento de dignidad personal. Estos rasgos de nuestra sociedad lo hacen a uno dudar de que los aumentos modestos en las tasas marginales de impuesto al ingreso habrán de reducir el trabajo y, con él, los rendimientos fiscales, aun cuando el costo de recaudar los impuestos probablemente se elevará, al menos en una medida discreta, dado que un incremento fiscal aumenta el beneficio cosechado por la evasión fiscal.

El impuesto al ingreso incluye de más y de menos desde un punto de vista económico. Incluye de más porque grava un ingreso que el receptor ahorra en lugar de gastarlo en consumo. El dinero ahorrado genera ingreso futuro, el que también se grava, de modo que el mismo ingreso se grava dos veces. Supongamos que me encuentro en el estrato de 20% del grupo del impuesto al ingreso marginal y que estoy tratando de decidir entre consumir 100 dólares de mi ingreso o invertirlo en un bono que vencerá al cabo de un año y rendirá un interés de 5% (supongamos una inflación nula). Después del impuesto tendré 80 dólares para consumir o invertir. Si consumo, obtendré un beneficio neto medido por el excedente del consumidor que generarán los bienes y servicios que compro por 80 dólares, y supongamos que esa cifra es de 10 dólares. No pagaré ningún impuesto al ingreso por este beneficio. Pero si en lugar de eso invierto en el bono, pagaré un impuesto de 20% sobre los 5 de intereses que el bono me rinde y así sólo me quedaré con 4. Los servicios generados por el consumo y los servicios generados por el ahorro están así cargados diferentemente, lo que es ineficiente. No obstante, a causa del diferencial en el tratamiento del consumo y el ahorro, los costos del ahorro a menudo se ven contrarrestados por los beneficios, que incluyen la liquidez (si los ahorros se invierten en forma líquida) y el consumo fluido a lo largo del ciclo vital: por ejemplo, el ahorrar en la juventud a fin de contar con efectivo para el consumo en la edad de la jubilación, cuando los ingresos bajan.

El ahorro de una parte de los ingresos derivados del trabajo responde de forma característica al propósito de permitirle a uno gastar en el consumo futuro. Esto hace que el impuesto al ingreso sea en parte un impuesto a la venta en el futuro consumo, un impuesto a añadir al impuesto sobre la venta que uno tendrá que pagar cuando dicho consumo tenga lugar. Expresado de otra manera, penaliza arbitrariamente el consumo futuro.

La preocupación por la doble tributación es uno de los argumentos a favor de la sustitución del impuesto al ingreso por un impuesto al consumo, que podría tener diversas formas, tales como un impuesto al valor agregado (que se discute en la siguiente sección) o un impuesto al flujo de efectivo con una deducción por el ingreso que se usó en la inversión antes que en el consumo.[32]

[32] El argumento a favor de un impuesto al gasto o al consumo se defiende enérgicamente en William D. Andrews, "A Consumption-Type or Cash Flow Personal Income Tax", 87 *Harv. L. Rev.,* 1113 (1974). Los problemas de la aplicación se discuten en Michael J. Graetz, "Implement-

Las objeciones se refieren principalmente a los efectos distributivos, y son de dos clases completamente diferentes. Una se concentra en la transición de un impuesto al ingreso a un impuesto al consumo. Con un impuesto al ingreso, las compras de bienes de capital de las empresas (bienes que generan un rendimiento durante un periodo futuro considerable, convencionalmente más de un año, antes que inmediatamente) se deprecian en lugar de deducirse; con un impuesto al consumo se deducen (de otro modo se gravaría el dinero empleado en la generación de ingreso futuro, junto con el ingreso futuro cuando se obtenga: ésta es la doble tributación que el impuesto al consumo trata de evitar). Supongamos que, antes del cambio del impuesto al ingreso al impuesto al consumo, una empresa hubiera comprado una planta, que planeaba depreciar, deduciendo la reserva de depreciación del ingreso que recibiera de los bienes vendidos por la planta. Después del cambio, todo el ingreso sería gravable, porque en el sistema de un impuesto al consumo son deducibles los gastos de capital cuando se hacen, en lugar de que sean deducibles bajo la forma de reservas de depreciación anual a lo largo de la vida del bien de capital.

La segunda y más convencional objeción distributiva a un impuesto al consumo consiste en que favorece a los ricos, quienes ahorran un porcentaje mayor de su ingreso.[33] Pero un impuesto al ingreso grava realmente sólo a una pequeña parte del ingreso proveniente del capital. La razón es que el impuesto reduce la varianza del rendimiento del contribuyente, dando por sentado que es adverso al riesgo. Supongamos que en cierto año vende el contribuyente acciones con una ganancia de 100 dólares y al año siguiente vende acciones con una pérdida de 100 dólares. Si se encuentra en el estrato del impuesto de 15%, su ganancia neta será de 85 dólares en el primer año y su pérdida neta será de 85 dólares en el segundo año. Al disminuir de esta manera el riesgo de su cartera, será inducido a aumentar el riesgo no diversificable de su cartera a fin de mantener su nivel de riesgo preferido, y en el capítulo XVI vimos que el riesgo no diversificable se correlaciona positivamente con el rendimiento; de manera que su ingreso se elevará. Si suponemos

ing a Progressive Consumption Tax", 92 *Harv. L. Rev.,* 1575 (1979), y en David F. Bradford, *Fundamental Issues in Consumption Taxation* (Am. Enterprise Inst. for Public Policy Research, 1996).

[33] Joseph Bankman y Barbara H. Fried, "Winners and Losers in the Shift to a Consumption Tax", 86 *Geo. L. J.,* 539 (1998); Alvin C. Warren, Jr., "Fairness and a Consumption-Type or Cash Flow Personal Income Tax", 88 *Harv. L. Rev.,* 931, 935 (1975). La mejor prueba de ello es que los que son ricos dejan legados mucho más grandes, en relación con su ingreso, que quienes no lo son. Las hipótesis alternativas en el sentido de que los legados son sencillamente una consecuencia fortuita de no vivir tanto tiempo como uno lo esperaba, o de que van encaminados a comprar servicios a cumplir por los futuros herederos de uno, resultan poco persuasivas. La débil demanda de rentas vitalicias, que darían garantía de que al morir una persona se quedara sin activos, sugiere que a los ricos los anima una fuerte motivación de dar un legado, y en una sociedad moderna los padres, particularmente los acaudalados, no dirigen la vista hacia sus hijos en busca de apoyo al llegar la vejez. (En el próximo capítulo se abordan más estos puntos.)

que los inversionistas ajustan sus carteras bajo esta modalidad, el impuesto al ingreso gravará en efecto sólo el componente libre de riesgo del rendimiento de su inversión.

XVIII.7. Impuesto al valor agregado

Como resultado parcial de los déficits presupuestarios en que incurrió la administración de George W. Bush, del hundimiento en los beneficios federales derivados del impuesto al ingreso merced a la severa desaceleración económica iniciada en 2008, del sostenido crecimiento de los gastos de derechos conforme la población nacional envejece y de los nuevos programas de gastos, la deuda nacional se ha vuelto estratosférica. Debido a que el presidente Obama ha prometido no elevar las tasas de impuestos al ingreso aplicados a las personas que ganan menos de 250 000 dólares anuales, así como a que los grupos de interés bloquean los recortes de mayor proporción al gasto federal, se está sopesando nuevamente (si bien no por parte del Congreso) la conveniencia de establecer un impuesto al valor agregado federal: una forma de impuesto al consumo. Un impuesto de esa índole genera rendimientos cuantiosos a un costo bajo por la recaudación, y opera como un impuesto al consumo, alentando de esa forma la inversión. También es cierto que incrementa la desigualdad de ingresos, pero su rendimiento potencial es de tal naturaleza que los gastos federales habidos por la redistribución de la riqueza y los beneficios correspondientes al bienestar social podrían aumentar notablemente.

En la clase más simple de impuesto al valor agregado, el minorista paga el impuesto, que es un porcentaje de la diferencia entre el precio al que vende sus artículos o servicios y el costo de su mano de obra, sus materiales y otros gastos comerciales (tales como la renta y el seguro). En otras palabras, está pagando el impuesto al valor que añadió a los insumos que ha comprado. De modo semejante, sus suministradores pagan el IVA fijado sobre los beneficios de sus ventas menos sus costos, esto es, sobre el valor que *ellos* agregaron en el proceso de producción-distribución.

Ya que (sin suponer excepciones) la base fiscal para el IVA es tan amplia (todos los bienes y servicios), un IVA puede generar inmensos rendimientos fiscales a una tasa tributaria baja, lo que reduce el efecto distorsionador del impuesto. Además, un impuesto proporcional al valor agregado es un representante razonable aunque elemental de los servicios gubernamentales que recibe una empresa (y, por ello, sus clientes). El IVA también evita la doble tributación de los ahorros con un sistema de impuesto al ingreso de tipo corporativo más individual, y asimismo fomenta el ahorro al hacer el consumo más costoso, así como reduce los efectos desincentivadores de la tributación onerosa al ingreso. El IVA ha sido el blanco de críticas que lo tachan de regresivo, pero si la tasa fiscal es baja el efecto regresivo se aligera, y en cualquier caso podría ser contrarrestado por subsidios a los pobres. También se critica al

IVA por juzgarlo inflacionario, en virtud de que determina un aumento en los precios de menudeo y, por lo tanto, en los índices de precios al consumidor. Con todo, las pruebas empíricas (la mayoría de las naciones poseen un IVA, así que existen abundantes datos para examinar sus efectos) señalan que, en el peor escenario, lo único que ocasiona es un problema minúsculo por una sola vez en el índice.

Por supuesto, los beneficios del IVA son mayúsculos si viene a remplazar a los impuestos sobre la renta y otros impuestos ineficientes en lugar de ser añadido al sistema tributario existente con el propósito de generar rendimientos fiscales adicionales. (El efecto inflacionario desaparecería si el IVA meramente sustituyera a otros impuestos que se les trasladan a los consumidores.) No obstante, lo que sucede es que no solamente tal sustitución es poco factible por motivos políticos, sino que el costo resultante de transitar de un sistema tributario basado en el ingreso a uno basado en el valor agregado en sucesivos estadios de la producción y la distribución sería inmenso.

Los conservadores ponen objeciones al IVA porque, como ya se hizo notar, les ponen peros a cualesquiera impuestos que se asomen en el horizonte (como el impuesto a las ventas por Internet), incluso a los eficientes, y que no sólo remplacen los impuestos aplicados hasta la fecha. Fundan su preocupación en que, al incrementar los rendimientos del gobierno, el impuesto aumentará el tamaño del gobierno en relación con la economía privada. Empero, esta posición es difícil de sostener ante el espectro de nuestra elevada deuda pública (una consecuencia de los recortes fiscales aplicados por la administración de Bush, la crisis económica comenzada en el otoño de 2008 y la elevación en los gastos de derechos derivada en gran medida del número creciente de estadunidenses de la tercera edad), la cual merced a una mezcla de razones políticas y económicas difícilmente se logrará tener bajo control por medio de tasas fiscales más altas, recortes en el gasto (o la tolerancia requerida para adoptar nuevos programas de gastos) o una tasa de crecimiento económico lo suficientemente veloz para contrarrestar la tasa de crecimiento de dicha deuda. La inflación es la respuesta estándar a la deuda en exceso, pero infligiría un perjuicio a nuestro comercio exterior al volver nuestras exportaciones más costosas. Y la tasa inflacionaria tendría que ser peligrosamente alta, ya que la mayor parte de nuestra deuda pública es a corto plazo, de modo que la inflación moderada sería virtualmente inefectiva en su totalidad debido a que las tasas de interés sobre la deuda se ajustarían con rapidez. Las tasas de inflación que son elevadas producen efectos negativos sobre la economía.

Un efecto inmediatamente benéfico de adoptar el iva, a condición de que no se añadiera simplemente a los impuestos que ya existen, sería el de reducir los precios de exportación y, por consiguiente, aumentar las exportaciones, haciendo así que sobreviniera una disminución en el déficit comercial del país, el cual constituye uno de los factores que contribuyen a la deuda pública. El Acuerdo General sobre Aranceles Aduaneros y Comercio permite que el

IVA sea bonificado en las exportaciones, gracias a lo que disminuye el costo de los bienes exportados para sus compradores en el extranjero y así estimula la demanda de artículos estadunidenses en otros países.

El fantasma de John Maynard Keynes se aparecería para objetar el IVA porque gravar el consumo reduce la inversión, un tema conspicuo en *Teoría general del empleo, el interés y el dinero* (1936). Tendemos a creer que la verdad es lo contrario: que los ahorros financian la inversión y, entre más consuman las personas, menos ahorrarán. Pero, como señaló Keynes, las personas que ahorran y las personas que realizan inversiones empresariales son distintas y obedecen a propósitos distintos. Los ahorradores pueden no abrigar el deseo de financiar la inversión empresarial a causa de sus riesgos; tal fue ciertamente el caso en 2009, cuando la tasa de ahorro personal se elevó marcadamente, pero la inversión privada se desplomó. Las personas ahorraban más pero sus ahorros se mantenían inertes; adoptaban la forma de efectivo o de cuasi equivalentes de efectivo, como por ejemplo los bonos del Tesoro o los depósitos asegurados federalmente. La justificación para el gasto de estímulo en un escenario de depresión o de profunda recesión es que el gobierno, al tomar prestados ahorros pasivos y ponerlos en funcionamiento a fin de financiar proyectos públicos, encara el reto de reactivar una economía floja creada por la mengua en la inversión privada. (La Junta de la Reserva Federal respondió a la crisis económica parcialmente bajando las tasas de interés y, de este modo, inundando la economía de dinero; uno de sus propósitos era propiciar la inversión productiva haciendo que los rendimientos sobre las inversiones pasivas de carácter seguro tales como los bonos gubernamentales fuesen tan magros que incluso los inversionistas cautelosos desplazarían parte de su efectivo hacia las acciones, alentando así la actividad empresarial.)

El gasto por consumo, según nos enseñó Keynes, en realidad estimula la inversión, puesto que incrementa tanto los ingresos como la demanda de bienes y servicios, toda vez que el gasto constituye un ingreso para el vendedor del artículo sobre el que ha puesto la vista el consumidor, de quien a su vez puede esperarse que utilice parte de su ingreso para comprar algo más de algún otro comprador. Un impuesto al consumo reduciría la cantidad de consumo sin necesariamente aumentar la inversión, aunque si contribuyera en forma significativa a equilibrar el presupuesto federal reduciría las tasas de interés, lo que tendería a estimular el consumo al incrementar el préstamo.

XVIII.8. La definición del ingreso

La exclusión de cualquier tipo de ingreso del ingreso gravable reduce la base gravable, crea incentivos para trasladarse a actividades que generen la forma del ingreso excluida y, dependiendo de la riqueza de quienes reciben la forma del ingreso excluida, puede aumentar la desigualdad. Pero ¿cómo debe

entenderse el "ingreso"? La definición más amplia de ingreso sería el total de las entradas pecuniarias y no pecuniarias, no sólo del ocio y (otros) ingresos no pecuniarios derivados de la producción hogareña, sino también de las donaciones, las herencias y los premios. Si nuestra preocupación preeminente es la minimización de los efectos de sustitución del impuesto al ingreso personal, la interrogante de si deben incluirse o no en el ingreso gravable las donaciones, las herencias y los premios podría depender de que éstos sean donativos puros o, en efecto, compensaciones por el trabajo. ¿Es el regalo de un padre a un hijo una compensación por los servicios prestados o esperados, o una pura expresión de amor (altruismo)?[34] Si es lo primero, debería gravarse para evitar que los hijos se vieran inducidos a sustituir la producción para el mercado por la producción hogareña, en vista del trato fiscal diferente dispensado al ingreso proveniente de estos empleos alternativos. Si es lo último, ningún impuesto a la donación causaría sustitución alguna en contra del empleo en el mercado, mientras que gravarla sería una forma peculiarmente costosa de tributación, ya que reduciría la utilidad de tres personas (los padres y el hijo) a fin de recaudar dinero de una sola. Este examen sugiere una base analítica para tratar como ingreso a las donaciones que ocurren en el curso de una relación comercial, y (con menos vigor) para no tratar como ingreso a otras donaciones.[35]

A veces se argumenta que las donaciones deberían quedar excluidas del ingreso gravable porque no son parte del ingreso nacional, ya que al ser puros pagos de transferencia no aumentan la riqueza nacional. Esto es cierto… pero resulta irrelevante para la política fiscal. El propósito de la tributación consiste en recoger las recaudaciones para el gobierno, y un buen sistema tributario es aquel que lleva a cabo esta tarea con una cuota mínima de sustitución y de efectos distributivos (indeseables). Ya sea que la actividad gravada genere o no riqueza, no ocupa un lugar aquí ni allí.

Adviértase que el problema abordado en esta sección —la definición de ingreso— es idéntico al problema abordado en la primera sección del capítulo XVII —la medición de la desigualdad—. El problema radica en la diferencia entre el ingreso monetario y el ingreso real. Este último es un concepto mucho más amplio. Un impuesto limitado al ingreso monetario generaría efec-

[34] *Cf.* V.1 antes. Adviértase que no importa si la donación se realiza *inter vivos* (es decir, si tiene efecto antes de que sobrevenga la muerte del donador) o es testamentaria. Una donación testamentaria puede ser una recompensa por servicios desempeñados durante el curso de la vida del testador.

[35] Esta distinción se rechazó en Commissioner *vs.* Duberstein, 363 U.S., 278 (1960), pero fue finalmente restablecida por la promulgación del 274b del Código de Recaudación Interna vigente, que limita a 25 dólares la deducibilidad de la mayoría de las donaciones comerciales. Una donación que es una recompensa por los servicios prestados, por ejemplo, una donación a un conocido de la empresa o negocio que remitía clientes al donador, no es una donación en el sentido económico. Exentar a las donaciones comerciales del impuesto al ingreso meramente propiciaría que la compensación por los servicios prestados fuera designada una donación.

tos de sustitución muy indeseables, dado que hay muchos sustitutos para el ingreso monetario.

Aunque las donaciones motivadas por el altruismo crean un beneficio doble —aumentan la utilidad del donante y del donatario—, ése no es un argumento en contra de gravarlas. Por el contrario, entre mayores sean los beneficios de las donaciones, menos probable es que una tributación moderada de ellas tenga un efecto de asignación sustancial. Adicionalmente, algunos tipos de donativos son ineficientes y debieran ser desalentados porque abren paso a la búsqueda de renta por parte de donadores potenciales y al mismo tiempo, a través del efecto del ingreso, motiva que los donadores sustituyan el trabajo por el ocio, determinando una disminución en los ingresos fiscales.[36]

La ley fiscal solía distinguir entre los premios ganados en un concurso (gravables) y los otros premios tales como el Premio Nobel (no gravables). La distinción era poco sólida y ha sido eliminada. Si no se gravaran los premios de los concursos se induciría a la gente a sustituir con los concursos otras formas de la actividad productiva.[37] Y el hecho de no gravar a los ganadores del Premio Nobel y otros honores tendría un efecto similar. La existencia de estos premios afecta a las decisiones de investigación tomadas por personas en ocupaciones elegibles y aun la elección de ocupaciones. Aunque la elasticidad de la respuesta a los cambios ocurridos en el estatus fiscal (y, por ende, en los beneficios monetarios netos) de estos premios bien podría ser escasa, ése es un argumento a favor de que se los grave fuertemente en lugar de exentarlos del impuesto. Pero también debemos considerar la elasticidad de respuesta de las personas que aportan el dinero de esos premios a los cambios en el estatus fiscal de éstos. Un gravamen podría conducirlas a remplazar ese uso de su efectivo por otro que encerraría menos valor para ellas; éste sería un costo social del impuesto.

Los costos administrativos impiden una definición amplia de ingreso para fines fiscales, pero algunas de las exclusiones necesarias, en particular la del ocio, podrían tener efectos de sustitución considerables. En el caso del ocio la exclusión sesga la elección de actividades a favor de aquellas que generan ocio en lugar de ingreso pecuniario: actividades tales como la enseñanza, en las que las vacaciones largas pagadas[38] forman una parte impor-

[36] Véase Louis Kaplow, "Tax Policy and Gifts", 88 *Am. Econ. Rev. Papers & Proceedings,* 283 (mayo de 1998).

[37] Una lotería o concurso es una actividad productiva. Crea utilidad.

[38] Así pues, ¿qué significa "vacación pagada" para un economista? Cada vacación es parte del ingreso real de un individuo y, por ello, es una vacación pagada, ya sea que el pago de salarios continúe o se suspenda durante el periodo vacacional. El ingreso real de un maestro consiste en un salario más tres meses de vacaciones; no tiene relevancia económica que el salario se pague solamente durante los nueve meses de trabajo o a lo largo del año entero. Si las vacaciones de los maestros fuesen más cortas el salario pecuniario de la actividad docente sería más alto, dado que se le impondría al maestro un costo —el valor perdido del ocio para él— por el acortamiento de su vacación. Del mismo modo, entre dos trabajos que proporcionan el mismo ingreso real,

tante de la compensación del empleo. Se introduce un sesgo similar por la exclusión necesaria de la fama, el prestigio, la comodidad, lo emocionante y otros rendimientos de la actividad intangibles, así como por el hecho de que no se excluya la parte del ingreso que representa una compensación de los peligros o los aspectos desagradables de algunas ocupaciones (incluido el riesgo en el sentido económico). Pero muchas de las exclusiones no pueden justificarse por lo aplastante de los costos de medición. He aquí cuatro ejemplos: los servicios que no se venden en un mercado, la renta imputada de las instalaciones ocupadas por el propietario, las prestaciones accesorias y los daños cuasidelictivos.

1. A menudo hay una elección entre contratar un servicio en el mercado y producirlo en el hogar. Una ley del impuesto al ingreso que cuente como ingreso sólo las recepciones pecuniarias sesga la elección a favor de la producción hogareña. Es probable que el más importante de tales sesgos creados por la ley del impuesto al ingreso vigente surja del hecho de que no se incluya el considerable ingreso real (pero no pecuniario) generado por los servicios de las amas de casa dentro del hogar. Supongamos que una mujer pudiera ganar 10 000 dólares fuera de la casa, sobre lo que pagaría un impuesto al ingreso de 2 000, mientras que si permaneciera en su casa sus servicios valdrían sólo 9 000 dólares. El valor de su trabajo aumentaría si trabajara afuera, pero el efecto de la definición incompleta de la ley del impuesto al ingreso es inducirla a permanecer en el hogar. Ese efecto es particularmente significativo para los cónyuges que presentan una declaración de ingresos conjunta, porque el ingreso que gana la esposa trabajando fuera del hogar está gravado a la tasa marginal de su esposo y porque a menudo será elevada la elasticidad cruzada del trabajo hogareño y el trabajo en el mercado; es decir, la producción hogareña de la esposa, sobre todo si la pareja tiene hijos, podría ser tan valiosa para la pareja que sería un sustituto excelente para el trabajo fuera de la casa; en tal caso, un impuesto a los ingresos externos, pero no a los internos, tendrá considerables efectos de asignación. Y recuérdese que la meta del diseño de los impuestos es la minimización de tales efectos, o sea, interferir en la menor medida posible con la asignación de los recursos que existiría si no hubiese impuestos.

Consideraciones como éstas condujeron a la larga a la eliminación reciente del tope al ingreso ganado por los receptores de la seguridad social. Antes de esta reforma, una persona que fuese por lo demás elegible para los beneficios de pensiones de la seguridad social no tenía derecho a ellos entre

salvo que en uno el tiempo de ocio del trabajador es más prolongado y su ingreso monetario más bajo que en el otro, quien tenga el gravamen más pesado será el trabajador del otro empleo. Esta disparidad, que es ineficiente debido a que incrementa el costo privado más allá de su costo social, se deriva del hecho de que no se realiza ningún esfuerzo para gravar el ingreso bajo la forma del ocio.

las edades de 65 y 70 años si también tuviera ingreso ganado. Dado que el impuesto recaía de manera especialmente pesada sobre las personas de escaso potencial para obtener ingresos, la eliminación del tope redujo la desigualdad económica al mismo tiempo que eliminaba un poderoso desincentivo para desempeñar un empleo pagado.

Hay grandes dificultades para la valuación de los servicios que se prestan fuera del mercado; señalamos antes el error de equiparar el valor de los servicios de un ama de casa a los salarios de una empleada doméstica (véase antes VI.11). Pero, dado que tales salarios representan una estimación mínima del valor de la mayoría de los servicios de las amas de casa, incluir esa cantidad como ingreso imputado de cada familia en la que la esposa no trabajara fuera del hogar sería un paso en la dirección correcta. La ley fiscal ha dado algunos pasos vacilantes para reducir los efectos de sustitución causados por la incapacidad de gravar el ingreso imputado de las amas de casa otorgando créditos fiscales para los gastos moderados en la atención de los hijos[39] y para alguno de los ingresos obtenidos en el mercado por el cónyuge que es el contribuyente principal. En virtud de que los créditos son muy moderados y los créditos fiscales son muy progresivos por comparación con las deducciones, que son regresivas,[40] tienen escaso efecto sobre los incentivos de las mujeres que serían más productivas en el mercado. El crédito fiscal por la atención de los hijos es realmente muy extraño: estimula la entrada al mercado de mujeres que son precisamente las que se ocupan más productivamente en el hogar, es decir, las madres de hijos jóvenes.

2. Un hombre da a sus dos hijos, A y B, 10 000 dólares a cada uno. A deposita sus 10 000 dólares en un banco de ahorros que paga 5% anual de interés, el que A utiliza para pagar la renta de su departamento. En virtud de que se encuentra en el estrato de 20% del impuesto federal al ingreso, A paga al gobierno 100 dólares de los intereses que recibe cada año. B, quien se encuentra en el mismo estrato del impuesto que A, utiliza sus 10 000 dólares para comprar un departamento que tiene el mismo valor de renta que el departamento rentado por A. B no paga ningún impuesto al ingreso, aunque ha destinado la misma cantidad de dinero al mismo uso que A. Simplemente, B está en mejor situación que A por la cantidad de 100 dólares anuales. Esta diferencia arbitraria del tratamiento genera un incentivo (incrementado por la deducción de los intereses) para que la gente sea propietaria en lugar de

[39] William D. Popkin, "Household Services and Child Care in the Income Tax and Social Security Laws", 50 *Ind. L. J.,* 238 (1975).

[40] Por lo común, las deducciones son proporcionales al ingreso, de manera que los contribuyentes de mayores ingresos se salvan de pagar una cantidad más elevada de impuestos. Los créditos fiscales, que usualmente se fijan en cantidades en dólares en vez de establecerse en porcentajes, son progresivos porque producen un mayor porcentaje de reducción de los gravámenes entre más bajos sean los impuestos que debe el contribuyente.

rentar sus casas, y para que los abogados creen complejas formas legales, tales como los condominios, destinadas a permitir que las propiedades rentadas aparezcan como propiedades simples de pago.[41] No obstante, la ventaja fiscal, hasta donde puede anticiparse, será capitalizada en el precio del condominio, ya que al decidir cuánto ha de pagar por una casa el consumidor racional toma en consideración todos los costos asociados a ella, como los impuestos, y de este modo pagará más entre más bajos sean los gravámenes.

Habría dificultades administrativas para estimar el ingreso de renta real pero no pecuniario que obtiene la gente del hecho de ser propietaria de sus casas en lugar de rentarlas a otras personas. Pero aun una estimación burda, sesgada hacia abajo, reduciría el incentivo para sustituir la casa rentada por la casa propia. Obsérvese que, una vez dado este paso, la deducibilidad de los pagos de intereses sobre la hipoteca de la casa, criticada previamente en este capítulo, dejaría de ser objetable, porque los pagos serían entonces un gasto para la producción de ingreso gravable.[42]

3. La exclusión de las prestaciones del ingreso gravable otorga a los ejecutivos un incentivo para tomar una parte de su compensación bajo la forma de una oficina elegante, un automóvil de la compañía, etc., aunque estas cosas valgan menos para ellos (excluyendo las consideraciones fiscales) que su equivalente en efectivo; la diferencia es puro desperdicio. La exclusión de las prestaciones complementarias da a los empleados un incentivo para tomar una parte de su compensación bajo la forma de seguros de vida y de salud, vacaciones y beneficios de pensiones,[43] aunque prefirieran —aparte de las consideraciones fiscales— el equivalente en efectivo.

En principio, una prestación complementaria debiera gravarse para el empleado como la diferencia existente entre el costo del beneficio para el empleador y su valor para el empleado.[44] Supongamos que el empleador rembolsa al empleado por el uso de la guardería diurna que le cuesta al empleador 100 dólares por empleado, pero el valor de la guardería para el empleador (por cuanto reduce la cantidad de tiempo que un empleado debe tomar de su

[41] Henry Hansmann, "Condominium and Cooperative Housing: Transactional Efficiency, Tax Subsidies, and Tenure Choice", 20 *J. Legal Stud.,* 25 (1991). Inglaterra gravó alguna vez el ingreso de renta imputado. Richard B. Goode, *The Individual Income Tax,* 117-125 (ed. rev., 1976).

[42] En la siguiente sección se discute la deducción de los gastos en que se incurre en la producción de ingreso.

[43] En virtud de que los beneficios de pensión se gravan al ser recibidos, el participante en un plan de pensión tiene la facultad de desplazar el ingreso de un periodo durante el que está empleado y, por lo tanto, cumple los requisitos para ser clasificado en el grupo de los individuos sujetos a un impuesto al ingreso más alto en dirección hacia el periodo en el que se encuentra ya retirado y, presumiblemente, habrá de quedar clasificado en un grupo inferior.

[44] Véase Avery Katz y Gregory Mankiw, "How Should Fringe Benefits Be Taxed?", 38 *Nat'l. Tax J.,* 37 (1985). Véase también David A. Weisbach, "Ironing Out the Flat Rate", 52 *Stan. L. Rev.,* 599 (2000).

trabajo para atender a su hijo) es de sólo 60 dólares. Esto implica que la única razón por la que el empleador proporcionó la guardería de 100 dólares, en lugar de algún arreglo que le costaría sólo 60, era la de otorgar al empleado un ingreso adicional de 40 dólares que no fuera gravado.

Pero los costos no se ubican todos de un solo lado. Además de las dificultades de medición, la tributación de las prestaciones accesorias y las concesiones adicionales agravaría la tendencia del impuesto al ingreso a causar una sustitución del trabajo por el ocio, porque las prestaciones accesorias y las concesiones adicionales se correlacionan con el trabajo antes que con el ocio.

4. De acuerdo con la ley vigente, el pago de daños otorgado en un caso de perjuicio personal no es gravable aunque remplace a ingresos que sí lo son. Es muy dudoso que el público contribuyente se beneficie con esta exclusión.[45] El efecto consiste en aumentar el valor de los fallos en los juicios de perjuicios personales. Por supuesto, sin embargo, sólo una fracción minúscula de los contribuyentes recibirá alguna vez tales fallos. Entonces, si se derogara la exclusión, en igualdad de las demás circunstancias, la tasa del impuesto al ingreso sería ligeramente menor y todos los contribuyentes se beneficiarían. En efecto, el beneficio probabilístico de un pago de daños mayor se convertiría en una reducción impositiva actuarialmente equivalente, pero segura. Si suponemos la aversión al riesgo, esta conversión aumentaría el bienestar.

XVIII.9. Deducciones del impuesto al ingreso

La suposición debiera ser (y es) en contra de que se permita al contribuyente deducir los gastos incurridos en el cálculo de su ingreso gravable, porque un gasto no es generalmente una reducción del bienestar sino el paso necesario por el que la gente transforma el dinero en un aumento del bienestar: eres más rico, no más pobre, después de que has cambiado efectivo por un televisor nuevo; no habrías hecho el intercambio hasta que hubieras pensado que te colocaría en una mejor situación. Pero los gastos destinados sólo a producir ingreso no tienen este carácter. Tales gastos aumentan el bienestar sólo en la medida en que generen ingreso; por lo tanto, si el ingreso es plenamente gravable, la deducción de los gastos mercantiles es necesaria para evitar la doble tributación; si ganas 20 dólares y se te fija un gravamen a la tasa de 50% y no se te permite deducir los 10 dólares en que incurriste para ganar los 20, entonces la tasa de impuesto efectivo sobre tu ingreso neto es de 100%, no de 50%. Sin embargo, algunos gastos mercantiles son también gastos de consumo. Las comidas en el restaurante y el alojamiento en el hotel que compro en un

[45] Véase Thomas D. Griffith, "Should 'Tax Norms' Be Abandoned? Rethinking Tax Policy Analysis and the Taxation of Personal Injury Recoveries", 1993 *Wis. L. Rev.,* 1115 (1993).

viaje de negocios son gastos indispensables para la producción del ingreso que obtenemos mi empleador o yo del viaje, pero al mismo tiempo son actividades de consumo en las que me ocuparía en alguna forma si me hubiera quedado en casa.

La ley se ha ocupado de este problema permitiendo la deducción de todos los gastos de subsistencia (a menos que sean "dispendiosos") en que se incurre en los viajes de negocios de un día para otro, al mismo tiempo que prohíbe como gastos "personales" la deducción de cualquier gasto para ir al trabajo y la mayor parte de otros gastos vitales en los que se incurre en el lugar donde tenemos la oficina. Esto podría parecer arbitrario, pero alguna distinción entre los gastos de viaje locales y foráneos tiene sentido económico. Aun si no trabajáramos, tendríamos que comer; dado que el costo de comer no se puede evitar dejando de trabajar, no es un costo del trabajo. El transporte suburbano permite que un trabajador viva en el lugar que le guste antes que al lado de su lugar de trabajo; por lo tanto, es tanto un gasto personal como de negocios. Por supuesto, hay algo de consumo personal en el viaje suburbano, pero es probable que su valor sea menor que el del gasto total, y la diferencia es un gasto puramente mercantil.

En principio, la diferencia, y no todo el gasto, es lo que debiera ser deducible del ingreso gravable.[46] Supongamos que realizo un viaje de negocios de un día para otro a San Francisco, me hospedo en un hotel que cobra 100 dólares diarios, y consumo alimentos que me cuestan un total de 150. La suma total de 250 dólares no debe ser deducible; el ahorro de gastos personales en consumo posibilitado por el viaje es lo que debiera restarse. El componente del alojamiento de lo que se resta será de ordinario pequeño, y por conveniencia administrativa es posible que debiera igualarse a cero, porque el costo marginal (limpieza, depreciación, electricidad, etc.) de pasar una noche más en casa es trivial. Lo que se resta en los alimentos será mayor, pero menor de 150 dólares, porque el costo de los alimentos preparados en casa es generalmente menor que el costo de las comidas en restaurantes. Es cierto que una comida de 100 dólares en un restaurante vale más para los comensales que la comida de 10 dólares que habrían disfrutado en casa; de otro modo, se habrían quedado en casa. Pero, dado que muy pocas personas comen frecuentemente en restaurantes caros, es probable que los 100 dólares que gasté en comidas fuera de casa en mi ejemplo hipotético me rindieron bastante menos que 90 dólares de valor medido por la disposición a pagar. Quizás el uso de una cifra intermedia, tal como 50% del precio de los alimentos de negocios, fuese un procedimiento apropiado para implantar la reforma sugerida.[47] Empero, una complicación es que resulta fácil subestimar el costo de la alimentación

[46] El gobierno aplicó este enfoque hace mucho tiempo, pero hubo de abandonarlo debido a su complejidad administrativa. La historia se narra en United States *vs.* Correll, 389 U.S., 299, 301 n. 6 (1967). Véase también Moss *vs*. Commissioner of Internal Revenue, 758 F.2d, 211 (7th Cir., 1985).

[47] El Congreso ha hecho enmiendas al Código de Recaudación Interna a fin de permitir que

en casa excluyendo el costo laboral de la persona que la prepara, que no es pecuniario pero sí real.

La única diferencia entre las prestaciones accesorias parece plantearse entre excluirlas del ingreso gravable y deducirlas del ingreso gravable. Pero en el primer caso hay un gasto del empleador, quien tiene por lo tanto un incentivo para economizar. En el último caso el empleador no incurre en ningún costo, y esto aumenta la elasticidad cruzada entre el consumo ordinario financiado por el ingreso después de impuestos y el consumo relacionado con los negocios financiado por el ingreso antes de impuestos.

Las deducciones de gastos médicos y debidas a pérdida por siniestro, percance o desastre que aparecen en el código fiscal parecen estar motivadas por un esfuerzo para distinguir entre los gastos que aumentan el bienestar personal y los gastos destinados sólo a restablecer al contribuyente en un estado de bienestar anterior.[48] Se trata supuestamente de igualar las cargas fiscales entre individuos que podrían tener ingresos pecuniarios idénticos pero ingresos reales diferentes. Podemos simpatizar con la meta, pero cuestionar su viabilidad y los medios específicos adoptados para alcanzarla. Consideremos la prima salarial que se paga a personas empleadas en una ocupación peligrosa. La prima salarial no aumenta su bienestar por comparación con el de las personas que tienen empleos nada riesgosos; sólo las compensa por el peligro. Pero sería inviable permitir que la prima se dedujera, puesto que resultaría difícil determinar cuánto del salario de un empleado era una compensación por la peligrosidad del trabajo. O consideremos a dos familias de ingreso igual, una más eficiente en el consumo que la otra, de modo que obtiene un bienestar mayor de la misma cantidad de dinero; también en este caso sería inviable una deducción para igualar el bienestar después de impuestos de las dos familias. En estos y muchos otros sentidos el ingreso pecuniario relativo distorsiona el bienestar, y, dado que la mayoría de estas distorsiones no pueden ser corregidas, resulta dudoso que un esfuerzo por corregir una o dos de ellas contribuya mucho a reducir la desigualdad económica.

Además, las deducciones específicas están mal diseñadas en relación con la meta. Cabe sostener que la medida de la deducción médica debiera ser el costo de la enfermedad para el contribuyente antes que el costo del tratamiento médico. Algunas enfermedades pueden ser muy costosas en términos de pérdida de utilidad a causa del dolor, el sufrimiento, la desfiguración u otras disminuciones del bienestar, pero, debido al nivel del avance en el arte médico o a la situación financiera del contribuyente, es posible que no se traduzcan en gastos sustanciales en el tratamiento. No obstante, en la pro-

se deduzca sólo 50% del precio de las comidas de oficina a menos que se trate de un viaje nocturno, caso en el que es deducible el precio total. Véase 26 U.S.C., 274(n).

[48] Se aboga por la distinción en William D. Andrews, "Personal Deductions in an Ideal Income Tax", 86 *Harv. L. Rev.,* 309 (1972), y Richard A. Epstein, "The Consumption and Loss of Personal Property under the Internal Revenue Code", 23 *Stan. L. Rev.,* 454 (1971).

porción en que la enfermedad reduce el ingreso del contribuyente, también reduce su impuesto al ingreso, de manera que obtiene un beneficio tributario indirecto de la enfermedad aun sin tener la facultad de deducir su costo directamente.

Y a veces se incurre en gastos médicos a fin de aumentar el bienestar (por ejemplo, una operación para desaparecer las arrugas o enderezar la nariz) y no para tratar una enfermedad u otra privación del bienestar; además, estos ejemplos sugieren cuán tenue es la distinción existente entre el aumento y la restauración del bienestar. Sin embargo, una parte de la deducción médica podría ser justificable en términos económicos como una reparación de la inversión en el capital humano. Como sabemos, el capital humano es tanto un activo como lo es una fábrica; se compra con insumos de tiempo y de dinero y rinde un servicio pecuniario durante la vida del activo (a veces de muy larga duración). Dado que el ingreso que genera es gravable, los costos de su mantenimiento en buen estado debieran ser deducibles, aunque en principio no de un solo golpe sino a lo largo del periodo en el que los gastos generan ingresos. Puesto que, si los gastos pudiesen ser deducidos antes de recibir el ingreso que han generado, esto significaría que son deducidos de un ingreso al que no han contribuido en forma alguna a generar. El asunto no importaría si el ingreso neto fuese siempre gravado a la misma tasa; no obstante, dado el carácter progresivo de la tributación al ingreso, los gastos a ser deducidos antes de haber recibido el ingreso que generaran le permitirían a la gente asignar los gastos a los años marcados por una tributación más onerosa.

¿Significa ello que, en lugar de tener simplemente una deducción médica, el costo total del capital humano debiera amortizarse a lo largo de la vida laboral del propietario y ser deducible de su ingreso? Probablemente no. Gran parte del costo de la educación, que es la fuente principal del capital humano junto con nuestra dotación natural de inteligencia, energía y carácter —una dotación que el propietario no paga tampoco—, no es soportada por el propietario. Y el costo mayor de la educación para el estudiante —los ingresos que se pierden por asistir a la escuela en lugar de trabajar— se paga en efecto en el momento en que se hace la inversión, ya que el ingreso gravable se reduce por el monto total de los ingresos sacrificados. No hay necesidad de una deducción adicional, sobre todo en vista de que la educación es un insumo de la utilidad no gravada, al igual que del ingreso.[49]

[49] En lo que respecta al tratamiento fiscal económicamente sólido del capital humano que se genera a través de la capacitación en el puesto de trabajo: el empleador puede deducir los gastos por la capacitación, aun cuando, en principio, éstos se deberían amortizar a lo largo de la vida laboral del sujeto dentro de la empresa. Entre más general sea el capital, más probable resulta que acabe por adscribirse al trabajador en lugar de a la empresa, y por ello el costo lo ha de sobrellevar aquél bajo la forma de un salario más reducido, al igual que sucede con la educación formal. Paul B. Stephan III, "Federal Income Taxation and Human Capital", 70 *Va. L. Rev.*, 1357 (1984). Para un panorama general, véase John A. Litwinski, "Human Capital Economics

En cuanto a la deducción por pérdida a causa de defunción, ésta se mide al menos por la pérdida antes que por el gasto en reparaciones. Con todo, su efecto no es tanto el de compensar a la gente cuyo bienestar ha sido objeto de merma como compensar a la gente que no se ha mostrado capaz de prevenir el evento comprando un seguro. La deducción reduce igualmente el costo de autoasegurarse, primordialmente en beneficio de los contribuyentes acaudalados. Un autoasegurador es alguien que prefiere asumir el riesgo de pérdida a pagar una cantidad fija (una prima de una compañía aseguradora) equivalente a la pérdida esperada. Si se pasa por alto el gasto administrativo, la pérdida esperada y el costo de las primas son los mismos, lo que hace de la pérdida esperada del autoasegurador el equivalente de las primas pagadas por el asegurado. De hecho, el gasto administrativo vuelve al seguro ofrecido en el mercado más costoso de lo que es el autoaseguramiento. Y ello por la razón adicional de que, si bien las primas de seguros no son deducibles de impuestos, la pérdida esperada (la contraparte de las primas que constituye el autoaseguramiento) es en efecto deducible, ya que en caso de la ocurrencia de una pérdida ésta será deducible bajo la provisión de pérdida por defunción.

Las deducciones médicas y por defunción también pueden ser criticadas con la misma base que la usada para la ausencia de tributación de los pagos de daños en los casos de lesión personal. Efectivamente, las deducciones les dan a las personas un seguro por el que no han pagado; y, si lo hubieran pagado, sus primas no habrían sido deducibles a partir de su ingreso gravable, del mismo modo en que los pagos de daños en los casos de cuasidelitos remplazan al ingreso que la víctima del cuasidelito habría obtenido de trabajar si no hubiera quedado discapacitada; ese ingreso lo habría obtenido aplicando su fuerza laboral y por ello habría sido gravable.

Una deducción importante, desde el punto de vista político y económico, es la deducción de las caridades. Es políticamente importante porque transfiere del gobierno al contribuyente algo del poder para decidir quiénes serán los receptores de las transferencias altruistas, una decisión que en la mayoría de las sociedades se toma a nivel político. Es económicamente importante porque responde al problema del gorroneo en la aportación caritativa que se mencionó en el capítulo XVII. La deducción por caridades puede ser en efecto más eficiente que la caridad directa del gobierno para inducir gastos en caridades. Si, como han descubierto algunos estudios empíricos, la elasticidad-precio de las caridades es mayor a 1, lo que significa que una reducción de 1% del precio conduce a un aumento mayor a 1% en la cantidad donada, la deducción de las caridades le cuesta al Tesoro menos, en recaudación perdida, que la ganancia de las caridades en aportaciones.[50] Si el tipo marginal del impuesto

and Income", 21 *Va. Tax Rev.,* 183 (2001); Jennifer J. S. Brooks, "Taxation and Human Capital", 13 *Am. J. Tax Pol'y,* 189 (1996).

[50] Martin Feldstein, "A Contribution to the Theory of Tax Expenditures: The Case of Charitable Giving", en *The Economics of Taxation,* 99 (Henry J. Aaron y Michael J. Boskin [comps.],

sobre la renta de las personas físicas (IRPF) fuera más bajo los costos de oportunidad de las donaciones de caridad se elevarían, pero este efecto bien podría ser contrarrestado, o incluso contrarrestado con creces, por el hecho de que la caridad es un bien superior (lo que significa que el monto de la caridad se eleva más que proporcionalmente con respecto al ingreso incrementado) y la reducción en la tasa fiscal aumentaría el ingreso posterior a la tributación de los contribuyentes que tienen ingresos elevados.

XVIII.10. El tratamiento especial de las ganancias de capital

La ley federal grava los ingresos sólo cuando se realizan.[51] Por lo tanto, si una corporación no distribuye todas sus ganancias entre los accionistas bajo la forma de dividendos, las ganancias no distribuidas no se gravan como ingreso personal. Las ganancias retenidas no son propiedad del accionista. Excepto en la liquidación, tales ganancias sólo lo enriquecen en la medida en que aumenten el valor de sus acciones. Cuando el accionista vende sus acciones, debe pagar un impuesto sobre la apreciación, incluida la parte de la apreciación imputable a la retención de ganancias. Pero esto no es un sustituto adecuado para el gravamen de la apreciación a medida que ocurre, porque posponer el impuesto otorga al contribuyente los intereses sobre la cantidad que habría pagado en impuestos si la apreciación se hubiera gravado cuando ocurrió.

El tratamiento fiscal favorable de la apreciación se fortalece cuando se gravan las ganancias de capital a una tasa inferior que la del ingreso ordinario…[52] lo que tendrá consecuencias ineficientes. Y consideremos la práctica irregular de las corporaciones de comprar sus propias acciones en el mercado abierto (esto es a veces, pero no siempre, también una estrategia contra la adquisición). El accionista que vende sus acciones a la corporación paga sólo un impuesto a las ganancias de capital; si en lugar de usar sus ganancias para comprar acciones la corporación hubiera pagado las ganancias en forma de dividendos, el mismo accionista habría tenido que pagar el impuesto regular al ingreso sobre la misma ganancia. La corporación podría obtener el mismo resultado reteniendo las ganancias y empleándolas en su negocio; la corpo-

1980); Jerald Schiff, *Charitable Giving and Government Policy: An Economic Analysis*, cap. 6 y p. 136 (1990).

[51] Quede el lector advertido de que los efectos de asignación, distribución y rendimientos de la tributación fijada a las ganancias de capital constituyen un área de inmensa incertidumbre en el campo teórico y en el empírico; el análisis de esta sección apenas si roza la superficie. Véase George R. Zodrow, "Economic Analysis of Capital Gains Taxation: Realizations, Revenues, Efficiency and Equity", 48 *Tax L. Rev.,* 419 (1993).

[52] Durante cierto tiempo las tasas eran las mismas, pero en la actualidad se ha restaurado el diferencial. La tasa fiscal máxima para las ganancias de capital a largo plazo (esto es, por más de un año) es hoy en día de sólo 15%, la tasa fiscal máxima para el ingreso ordinario es de 35 por ciento.

ración tendría más capital, y presumiblemente aumentaría el precio de sus acciones. Pero, dependiendo de la confianza que tengan los accionistas en la habilidad de los administradores para invertir estas ganancias rentablemente, podría no aumentar por el monto total de las ganancias retenidas. Dicho de otro modo, esas ganancias podrían generar un rendimiento mayor si fueran invertidas en otra parte. Ésta es probablemente la razón de que los fondos mutuos cerrados (donde el accionista no tiene derecho de amortización) venden de ordinario con un descuento sobre el valor de mercado de las acciones mantenidas por ellos. (Otra razón es que la inversión que realizan los inversionistas en las acciones subyacentes es sencillamente menos líquida.) Y recuérdese nuestra discusión de los dividendos como un método de reducción de los costos de agencia (antes, en XV.8). Sin embargo, la retención de las ganancias es un método alternativo para la explotación de una diferencia de las tasas impositivas entre las ganancias de capital y el ingreso ordinario, lo que significa que la diferencia de las tasas alienta lo que acabamos de decir que podría ser una práctica ineficiente. También induce a los inversionistas a sustituir con inversiones que rinden ganancias de capital, tales como las inmobiliarias, las inversiones igualmente productivas o más productivas que rinden ingreso ordinario.

Gravar las ganancias de capital tiene un efecto de "afianzamiento"; los propietarios de acciones se resisten a vender porque la venta es una realización de ganancia para propósitos fiscales. Por lo tanto, esta tributación reduce la eficiencia con la que el mercado de valores refleja los valores cambiantes del mercado.

Pero una cosa es *querer* gravar la apreciación no realizada cuando la corporación, en lugar de pagar las ganancias a los accionistas en dividendos, las retiene solamente a fin de que los accionistas puedan obtener esas ganancias bajo la forma de ganancias de capital (al vender su reserva en vías de apreciación), y otra es diseñar un sistema *práctico* para hacerlo. Cuando el activo de capital es una acción, resulta tentador gravar simplemente las ganancias retenidas de la corporación como ingreso de los accionistas, así como las ganancias que se pagan como dividendos. Pero, aparte del hecho fundamental de que el dinero sobre el que los accionistas no tienen un control directo no es realmente (o por lo menos no plenamente) su ingreso, la tributación de las ganancias retenidas plantearía formidables dificultades prácticas, sobre todo en el caso de las ganancias corporativas que debieran repartirse entre diferentes clases de valores.[53] Gravar la apreciación no realizada (por ejemplo, la

[53] Por ejemplo, entre los tenedores de pagarés convertibles (es decir, bonos convertibles en acciones) y de las acciones comunes. Las ganancias retenidas benefician a todos los tipos de tenedores de valores en una compañía, no solamente a los tenedores comunes. Incluso benefician a los tenedores de bonos, al incrementar el colchón de capital (lo que es uno de los motivos por los que los tenedores de bonos no son indiferentes a la política de dividendos de la compañía aun si la compañía no "invade el capital" al pagar dividendos). No obstante, la asignación de los

de los incrementos del valor de mercado de las acciones de la corporación, que podrían o no ser atribuibles a las ganancias retenidas) no plantearía estos problemas porque la apreciación es determinada por el mercado y sólo indirectamente por las ganancias de la empresa y por la forma en que éstas son asignadas entre las distintas clases de valores. De este modo, si las reservas de una compañía aumentan en valor, el incremento en la apreciación puede ser gravado sin hacer indagaciones acerca de la cantidad de ganancias retenidas. Pero ello plantearía otros problemas. A los contribuyentes les resultaría difícil pronosticar su responsabilidad fiscal. Los problemas de liquidez obligarían con frecuencia a los contribuyentes a vender valores u otra propiedad a fin de pagar impuestos sobre la apreciación no realizada. Las declaraciones fiscales resultarían más complicadas porque el gravamen fijado a las ganancias de capital no se limitaría a las transacciones. Y el impuesto corporativo tendría que ser replanteado en virtud de que los ingresos corporativos fluctuarían más de año en año a consecuencia de los cambios ocurridos en los precios de las acciones en el mercado, en los valores de los bienes raíces, etc., todo lo cual se vería reflejado en el ingreso.

Por último, la ventaja esencial de ser gravado sólo sobre la apreciación realizada, la cual consiste en que en efecto se permita al contribuyente acumular intereses libres de impuestos durante el periodo en que se difiera el pago, se contrarresta, por lo menos de una manera aproximada, por el hecho de que gran parte de la apreciación que se grava finalmente podría ser una apreciación fantasma causada por la inflación. Un ejemplo ilustrará esto. Supongamos que la tasa de interés real para alguna clase de valores moderadamente riesgosos es de 5%, la tasa de inflación es de 3%, y la tasa de interés del mercado es por lo tanto de 8%. Al cabo de 10 años, un bono que cuesta 10 000 dólares y que acumula intereses (compuestos anualmente) valdrá 21 589 dólares. Si se vende entonces y la tasa del impuesto a las ganancias de capital es de 25%, el impuesto ascenderá a 2 897 dólares. Descontada la inflación, la apreciación es de sólo 6 289 dólares, y después del impuesto es de sólo 3 392 dólares. Si no hubiera inflación,[54] la tasa de interés del mercado sería de 5%, el impuesto sería 1 572, dólares y la apreciación neta del tenedor de bonos después del impuesto sería de 4 717 dólares. Esto es más que en el caso de la inflación, lo que indica que algo de la apreciación gravada en ese caso era en efecto una apreciación fantasma causada por la inflación.

No obstante, al ser capaz de posponer el impuesto, el tenedor del bono fue capaz de ganar intereses sobre el ingreso que de otro modo habría tenido que pagar en impuestos. Si hubiera recibido pagos anuales por 800 dólares (8% de 10 000) sobre los que hubiese pagado un impuesto anual de 25%, su

beneficios de las ganancias retenidas entre las clases de tenedores de valores sería extremadamente difícil, casi imposible de lograr.

[54] Anticipada. Esta calificación es necesaria debido a que la inflación anticipada será incorporada a la tasa de interés, pero no ocurrirá lo mismo con la inflación no anticipada.

rendimiento después de impuestos habría sido de 6%, que compuesto a la misma tasa habría ascendido a 7 908 dólares para el final del periodo, por comparación con su rendimiento efectivo después de impuestos de 8 692 dólares [0.75 × (21 589 – 10 000 dólares)]. Entre más tiempo se difiera el impuesto, mayor será el monto de la tributación de la apreciación fantasma, pero al mismo tiempo será mayor el monto de los intereses ganados por el contribuyente a resultas de su capacidad para diferir el pago del impuesto.

Adviértase finalmente que el efecto de la tributación fijada a las ganancias de capital consiste en reducir el riesgo. El gravamen a las ganancias de capital trunca los rendimientos derivados de una venta del activo, especialmente si las pérdidas de capital pueden ser objeto de deducción, de modo que sea posible amortiguar tanto las pérdidas como las ganancias. Si los inversionistas son adversos al riesgo, el efecto de amortiguamiento les confiere una utilidad que contrarresta, parcialmente en todo caso, el costo del gravamen.

XVIII.11. El principio del impuesto progresivo

Un impuesto al ingreso es proporcional cuando todos pagan el mismo porcentaje de su ingreso. Una vez que la tasa impositiva se hace aumentar con el ingreso del contribuyente —se vuelve progresiva—, surgen varios problemas administrativos graves. Uno de ellos es el problema de la elección del momento apropiado. Con un impuesto proporcional, la persona que gana 10 000 dólares un año y 100 000 al año siguiente paga el mismo impuesto total que una persona que gane 55 000 dólares en ambos años. Con un impuesto progresivo al ingreso, la primera persona paga más impuestos que la segunda. Por lo tanto, se vuelve necesaria una disposición para promediar los ingresos para efectos del impuesto entre los años a fin de prevenir que la gente evite ingresar a ocupaciones en las que las ganancias varían desmedidamente de un año a otro. La Ley de la Reforma Tributaria de 1986 abolió la acción de promediar el impuesto al ingreso como parte de una reforma general que limitó el principio progresivo reduciendo la tasa impositiva marginal de los ingresos más altos a 28 por ciento.

La objeción más común a la tributación progresiva —más precisamente, a la elevada tasa impositiva marginal implicada por la tributación progresiva, porque, si el objetivo es que se obtengan abundantes rendimientos y a la vez se mantenga la progresividad, los grupos situados en lo alto tendrán que ubicarse en una posición muy elevada a fin de permitir que algunos de los grupos situados abajo tengan una posición inferior— es que provoca una sustitución ineficiente del trabajo por el ocio al aumentar el precio del trabajo (reduciendo equitativamente el ingreso neto derivado del trabajo) en relación con el del ocio. Pero si el ocio es un bien superior —si se consume proporcionalmente más ocio a medida que aumenta nuestro ingreso—, entonces las

elevadas tasas impositivas marginales asociadas a la tributación progresiva al ingreso podrían reducir el valor del ocio para el contribuyente de altos ingresos reduciendo su ingreso real (es decir, después del impuesto) en mayor medida de lo que reduce el costo del ocio para él en relación con el trabajo. Por cierto, esta observación revela que el aumento de la exención personal en el impuesto federal al ingreso reduce los incentivos para trabajar, además de ser regresivo. Es regresivo porque reduce el ingreso gravable, y por ello los impuestos, en mayor medida para quienes se encuentran en los estratos de elevado impuesto marginal que para quienes se encuentran en los estratos bajos. Reduce los incentivos del trabajo porque aumenta la riqueza del contribuyente sin reducir el precio del trabajo en relación con el del ocio, ya que la tasa impositiva marginal no se ve afectada.

Tal vez los efectos de ingreso deban ser pasados por alto, a la luz de nuestro examen previo de la manera en que la tributación al ingreso induce a los contribuyentes a aumentar la cantidad de riesgo sistemático en sus carteras y, por consiguiente, el ingreso proporcionado por las carteras (aun cuando debe hacerse una calificación, anotada más adelante). Aun así, si se omiten los efectos de ingreso, no está claro que un impuesto progresivo reduzca la cantidad de trabajo que realiza la gente. Porque si se mantiene constante el total de las recaudaciones fiscales, las tasas marginales del impuesto de algunos contribuyentes serán menores con un impuesto progresivo al ingreso de lo que serían con un impuesto proporcional. El efecto de esas tasas menores al aumentar los incentivos de trabajo podría ser mayor que el efecto de desincentivo de las tasas marginales mayores para otros contribuyentes. Con todo, si los contribuyentes de ingreso más alto son generalmente los trabajadores más productivos de la sociedad, el efecto de desincentivo del impuesto progresivo probablemente no se vería contrarrestado por el efecto de incentivo de las tasas marginales menores para los contribuyentes de menores ingresos. Ahora bien, los contribuyentes con ingresos elevados pueden ser muy productivos porque valoran el trabajo por encima del ocio, y, por consiguiente, pueden ser (dentro de ciertos límites) relativamente insensibles ante sus tasas impositivas marginales.

Esta referencia al efecto de la tributación progresiva sobre los incentivos para el trabajo podría ser un poco equívoca, ya que podría indicar aparentemente que el trabajo es más productivo que el ocio. Eso no es cierto en general. Lo importante es que no se asigne al trabajo un precio tan elevado por una tasa impositiva marginal elevada que se consuma una cantidad de ocio ineficiente.

El impuesto progresivo al ingreso podría reducir en efecto la cantidad de asunción del riesgo más allá de lo que ocurriría con un impuesto proporcional. Algunos ingresos elevados consisten en gran medida en la compensación del posible riesgo de terminar con un ingreso muy bajo. Supongamos que 1 de 10 compositores gana 100 000 dólares anuales por su actividad como compo-

sitor, mientras que los otros 9 no ganan nada. Aunque el ingreso medio del grupo es muy moderado, su ingreso se grava a una tasa mucho mayor que si cada compositor ganara 10 000 dólares.

Cuando se desalienta la asunción del riesgo se perpetúan las desigualdades del ingreso entre las familias. Es probable que no exista ninguna medida más eficiente para crear la movilidad del ingreso, hacia arriba y, por lo tanto, también hacia abajo, que la de alentar la asunción del riesgo. Los ricos que asumen riesgos y pierden acaban mucho menos ricos; los pobres que asumen riesgos y ganan acaban ocupando su lugar. Quizá sea por esta razón que la progresividad extrema no es muy popular entre los estadunidenses de medios moderados. Pero de nuevo debemos tener cuidado de extraer inferencias acerca de la eficiencia. Un impuesto proporcional al ingreso subsidia la asunción del riesgo, siempre que se permita que las pérdidas contrarresten el ingreso; porque en ese caso, como hemos visto, el impuesto reduce la varianza de los resultados de un proyecto riesgoso. Por lo tanto, las naciones que hacen relativamente escaso hincapié en la tributación del ingreso, *y* las naciones que cobran impuestos muy progresivos al ingreso, podrían ambas desalentar la asunción del riesgo y reducir así la movilidad social.

Un impuesto muy progresivo al ingreso otorga a los contribuyentes un incentivo para hacer fuertes gastos legales y contables con el propósito de evitar el impuesto. Si la tasa impositiva marginal fuese de 90% para el ingreso mayor a 50 000 dólares,[55] alguien con un ingreso de 200 000 dólares podría estar dispuesto a gastar hasta 135 000 dólares en medidas para no tener que pagar ningún impuesto en sus últimos 150 000 dólares de ingreso.[56] Los recursos consumidos a resultas de este gasto, como los costos en que se incurre en la monopolización —un caso paralelo—, serían improductivos. Cierto, el individuo de altos ingresos tendrá menos dinero para gastar con el fin de evitar los impuestos siempre que el sistema de tributación sea por lo menos parcialmente eficaz. Pero este efecto se contrarresta por el hecho de que los gastos para evitar el impuesto son deducibles de impuestos,[57] y también por el

[55] Hasta hace poco tiempo se encontraban tipos marginales del impuesto sobre la renta en este intervalo en algunos países europeos, tales como Suecia, y, entre la segunda Guerra Mundial y los primeros años del decenio de 1960 en los Estados Unidos.

[56] Esta conclusión no queda invalidada por el hecho de que a menudo se puede contratar a los abogados sobre la base de una cuota contingente a fin de obtener una reducción en la responsabilidad del contribuyente. En la medida en que el contribuyente puede pedir prestado dinero para financiar sus esfuerzos por evitar el gravamen, su falta de fondos personales no obstará para que contrate a un abogado con el propósito de que logre tener una respuesta favorable a su solicitud de una reducción de impuestos; el contrato por una cuota contingente es simplemente un método para pedir prestada una cantidad a fin de cubrir los gastos legales. Y es probable que el contribuyente, estando casi a punto de gastar los 135 000 dólares del ejemplo, dé un paso atrás, puesto que juzgará razonable proceder así cuando el (último) dólar marginal gastado compre sólo un dólar en impuestos reducidos.

[57] ¿Debieran serlo? Probablemente no en un sistema de tasas fiscales marginales de carácter elevado, en vista del incentivo que genera la deducibilidad para incurrir en gastos que son so-

hecho de que, dado que los gastos para evitar el impuesto producen ingreso (después de impuestos), un contribuyente podría ser capaz de obtener préstamos para financiarlos.

Los incentivos que las elevadas tasas marginales del impuesto generan para encontrar métodos que hagan posible evitar el impuesto podrían explicar el número de resquicios por los cuales pueden escapar ingeniosamente los contribuyentes de altos ingresos de la carga completa de un impuesto progresivo. Estas fisuras y estratagemas distorsionan los patrones del trabajo y la inversión. Las personas evitan las ocupaciones arriesgadas, como en el ejemplo anterior de los inversionistas, puesto que un ingreso progresivo de carácter elevado trunca los rendimientos indispensables para contrarrestar el riesgo. Las personas se agolpan en actividades tales como la de la inversión en bienes raíces en las que el ingreso se puede obtener bajo la forma de ganancias de capital, si es que la tasa de ganancias de capital es más baja que la tasa de ingreso ordinario. Y las actividades en las que el ocio es un componente importante de la remuneración, tales como la de la enseñanza universitaria, se vuelven más atractivas, dado que el ocio está exento del impuesto al ingreso. Los centros universitarios responden pagando salarios nominales más bajos, pero el efecto reside en habilitar a dichos centros para contratar personal docente a un costo más bajo que el afrontado por otros empleadores, lo que ocasiona una distorsión en la competencia entre los centros universitarios y otras empresas que encuentran más difícil remunerar a sus empleados con ocio. Las fisuras y estratagemas también frustran el propósito de la tributación progresiva de realizar la redistribución del ingreso de los más ricos a los menos ricos. Pero también reducen los ingresos reales de los primeros, debido no sólo a los gastos legales y contables en que incurren para minimizar los impuestos, sino también a los rendimientos inferiores generados por la gran parte del ingreso protegido contra los impuestos; la competencia entre los contribuyentes ricos ha empujado la tasa de interés de los bonos estatales y municipales libres de impuestos por debajo de la tasa de interés de las obligaciones de riesgo equivalente que no disfrutan de una exención fiscal. Asimismo, el individuo inducido por consideraciones fiscales a entregar una proporción mayor de su ingreso a obras caritativas, en relación con lo que haría de otro modo, obtiene del ingreso así gastado menos utilidad de la que habría obtenido en ausencia de elevados tipos marginales del impuesto sobre la renta. Por lo tanto, disminuye la desigualdad del ingreso, aunque quizá no se logre la redistribución de los ricos a los pobres. No está claro lo que se gana, a menos que haya mucha envidia en la sociedad.

El sistema del impuesto progresivo aumenta el atractivo político de la inflación como un instrumento de la política económica; porque, al empujar más

cialmente improductivos: gastos para crear y explotar resquicios en la redacción de los estatutos y regulaciones fiscales.

ingresos (o cualquier otra cosa que se grave progresivamente) a los estratos impositivos mayores, la inflación aumenta automáticamente las recaudaciones gubernamentales en *mayor* medida que el aumento de la inflación.[58] Pues si, en general, los ingresos se elevan de acuerdo con la tasa de inflación, un impuesto proporcional —esto es, un porcentaje plano— lograría cubrir adecuadamente al gobierno contra la inflación. Un impuesto progresivo al ingreso grava los ingresos nominales a una tasa de porcentaje más alta conforme se van elevando, de modo que el ingreso del gobierno se incrementa con mayor velocidad que la alcanzada por la tasa de inflación, y la inflación se convierte así en un método por el que el gobierno puede aumentar su ingreso real sin elevar la tasa impositiva o alterar de otro modo las leyes tributarias formales.[59]

Otro abuso al que la tributación progresiva se ha pensado que está peculiarmente sujeta es que supuestamente permite al electorado desplazar la carga de la tributación hacia una minoría numérica, integrada por personas de altos ingresos.[60] Pero, aparte de nuestra observación anterior acerca de la asunción del riesgo, la capacidad práctica de la mitad más pobre para desplazar la carga de la tributación hacia la mitad más rica está limitada por tres factores: los miembros de la mitad más pobre que esperan encontrarse algún día en la mitad más rica se sentirán renuentes a apoyar un impuesto al ingreso rápidamente progresivo, sobre todo porque la tributación del ingreso discrimina en contra de los nuevos ricos; la mitad de los más pobres podrían sentirse renuentes a adoptar un método de tributación que (por lo menos a cierto nivel de la progresividad) debe de perjudicar los incentivos para trabajar, ya que se benefician de la productividad de los ricos (o de los aspirantes a ricos) que son muy trabajadores, y las mayorías simples de los votantes no determinan con frecuencia la política pública (véase después XX.3).

Debe ser claro, por lo que vimos en el capítulo XVII, que la tributación progresiva no puede justificarse por referencia al principio de la utilidad marginal decreciente del ingreso. No sabemos si puede justificarse por la teoría del contrato de Rawls, como una medida destinada a maximizar el bienestar de quienes se encuentran en la peor situación. Concebiblemente, estos últimos estarían en mejor situación con una mezcla de la tributación proporcional —que debería estimular una actividad productiva mayor que la tributación progresiva y cuya administración podría ser más barata— y pa-

[58] Respondiendo a esta preocupación, el Congreso ha indexado ahora los estratos tributarios del impuesto al ingreso.

[59] Los impuestos a la propiedad no son progresivos, pero por supuesto aumentan en los periodos inflacionarios porque los valores de la propiedad —la base tributaria— están aumentando. No obstante, si el único impuesto fuera un impuesto sobre la propiedad, el efecto de la inflación sobre los ingresos tributarios no haría que la inflación constituyera una política agradable para los funcionarios de gobierno que desean ampliar el sector público porque los ingresos gubernamentales crecerían a la tasa de la inflación; con un impuesto progresivo, como se señaló en el texto, crecen más rápido.

[60] Milton Friedman, *Capitalism and Freedom,* 174-175 (1962).

gos de transferencia a los grupos de ingresos bajos. Aunque el trabajo no es más valioso que el ocio en un sentido económico, es más probable que produzca beneficios no captados por el actor; por ejemplo, el trabajo produce ingreso gravable y el ocio no lo produce, y las recaudaciones de los impuestos pueden utilizarse para ayudar a los pobres. Un impuesto proporcional podría también producir una mayor recaudación tributaria simplemente reduciendo los beneficios de las acciones para evitar el impuesto.

XVIII.12. Los subsidios fiscales

Un subsidio fiscal no es un gasto, tal como lo implica la palabra *subsidio,* sino una reducción fiscal selectiva, vista como algo diferente de una reducción general o uniforme. De ahí que eliminar un subsidio fiscal equivalga a elevar los impuestos. No obstante, lo de eliminar un "subsidio" fiscal suena como a reducir un gasto gubernamental que es dispendioso en vez de como algo encaminado a elevar los impuestos, de manera que goza de más atractivo popular que un incremento fiscal explícito. Pero no por ello se facilita la eliminación de un subsidio fiscal. El hecho de que los subsidios fiscales tiendan a tener en la mira determinadas actividades particulares significa que una propuesta de eliminar uno probablemente sirva de catalizador a la oposición formada por los grupos de interés, la cual configura una fuerza formidable, toda vez que si alguno de estos grupos fuera débil el subsidio fiscal no habría sido legislado en primer lugar.

Un subsidio es defendible con base en razones económicas si estimula la producción de beneficios sociales que de otro modo se habrían generado de forma deficiente. Tal es el argumento a favor de permitir que los gastos por investigación y desarrollo (I&R) se condonen (sean deducibles del ingreso gravable) en un proyecto o programa acelerado; la I&R se subproduce desde la perspectiva del bienestar social general debido a que, inclusive con un sistema de patentes, resulta bastante probable que la I&R de una empresa confiera beneficios a otras empresas por los cuales no se verá compensada. ¿Existe un caso similar para los subsidios fiscales específicos de determinada industria, por ejemplo, los subsidios para los productores de petróleo? Los principales subsidios fiscales al respecto son una "deducción a la fabricación interna" que les permite a las compañías petroleras y surtidoras de gas deducir un 6% extra de su ingreso gravable; una deducción por "costos intangibles", esto es, los costos por inversiones en la exploración o la producción de petróleo que no tienen valor de salvamento, tales como la del desmonte o limpia de tierras a fin de poner en marcha la perforación de un pozo petrolero (a las compañías petroleras no se les exige amortizar estos costos a lo largo de la vida entera esperada del pozo en cuestión) y, por último, se les permite a las compañías hacer una deducción por las regalías que paguen a los gobiernos extranjeros (la

teoría predica que las regalías pagadas a un gobierno constituyen en realidad un impuesto). Los valores agregados de estos subsidios a la industria petrolera estadunidense suman aproximadamente 5 000 millones de dólares al año.

Los primeros dos tipos de subsidios encierran la probabilidad de incrementar la producción petrolera interna, y la industria argumenta que la producción interna expandida crea beneficios externos al reducir nuestra dependencia respecto del petróleo extranjero, gran parte del cual lo producen naciones hostiles o inestables. (El tercer subsidio no puede ser defendido con este fundamento: incita a las compañías estadunidenses a incrementar su producción en el extranjero.) Esto es verdad, pero el efecto resulta probablemente pequeño, especialmente en relación con la práctica de imponer un arancel inamovible a las importaciones de petróleo. El arancel generaría rendimientos para el gobierno federal sin ser denominado un impuesto (si bien eso es lo que un arancel conforma), reduciría el ingreso de los países extranjeros productores de petróleo e incrementaría la producción interna en la medida en que hiciera más costoso el petróleo procedente de otras naciones.

Quienes abogan por eliminar los subsidios fiscales otorgados a las compañías petroleras sostienen que las ganancias de la industria petrolera son excesivas en relación con los elevados precios de la gasolina surtida en las gasolineras. No obstante, eliminar los subsidios se traduciría en precios más altos y no más bajos de la gasolina, puesto que de esa forma se reduciría la producción general de petróleo. Aun así, bien mirado, ¡ésa no sería una mala idea! Nuestros problemas con el petróleo no se limitan a las importaciones de éste, sino que incluyen el daño al ambiente (particularmente el efecto sobre el clima) causado por la quema y los derrames de petróleo, así como por el congestionamiento del tráfico. Los elevados precios de la gasolina, que reducen la demanda y por lo tanto el consumo, son el equivalente de un impuesto a la contaminación, y deberíamos fomentarlos.

Lecturas recomendadas

Alstott, Anne L., "Tax Policy and Feminism: Competing Goals and Institutional Choices", 96 *Colum. L. Rev.,* 2001 (1996).

Bankman, Joseph, y David A. Weisbach, "The Superiority of an Ideal Consumption Tax over an Ideal Income Tax", 58 *Stan. L. Rev.,* 1413 (2006).

Blum, Walter J., y Harry Kalven, Jr., *The Uneasy Case for Progressive Taxation* (1953).

Brennan, Geoffrey, y James M. Buchanan, *The Power to Tax: Analytical Foundations of a Fiscal Constitution* (1980).

Harberger, Arnold C., *Taxation and Welfare* (1978).

Institutional Foundations of Public Finance: Economic and Legal Perspectives (Alan L. Auerbach y Daniel N. Shaviro [comps.], 2008).

Kaplow, Louis, *The Theory of Taxation and Public Economics* (2008).
McCaffery, Edward J., "A New Understanding of Tax", 103 *Mich. L. Rev.,* 807 (2005).
Simons, Henry C., *Personal Income Taxation* (1938).
Staudt, Nancy C., "Taxing Housework", 84 *Geo. L. J.,* 1571 (1996).
Stephan III, Paul B., "Federal Income Taxation and Human Capital", 70 *Va. L. Rev.,* 1357 (1984).
Weisbach, David A., "An Efficiency Analysis of Line Drawing in the Tax Law", 29 *J. Legal Stud.,* 71 (2000).
———, "An Economic Analysis of Anti-Tax-Avoidance Doctrines", 4 *Am. Law & Econ. Rev.,* 88 (2002).
Zelinsky, Edward A., "Efficiency and Income Taxes: The Rehabilitation of Tax Incentives", 64 *Tex. L. Rev.,* 1973 (1986).

Problemas

1. ¿Puede imaginar algunas razones económicas a favor o en contra de las reglas tributarias siguientes?:

a) La separación del ingreso entre los esposos.
b) El tratamiento de un fideicomiso como una unidad tributaria.
c) La exención de los daños compensatorios en los casos de cuasidelitos.
d) Los impuestos del ingreso de mercado de una esposa a la tasa marginal del esposo.

2. Como vimos, una compañía podría ser capaz de desplazar un impuesto específico hacia delante, a los consumidores, y hacia atrás, a los proveedores. ¿Puede trasladarse el impuesto al ingreso personal bajo alguna circunstancia? Supongamos que se estableciera un recargo especial para el ingreso de los contadores. ¿Aumentarían los honorarios de los contadores? ¿Sería diferente su respuesta si los ingresos de los contadores incluyeran algunas ganancias monopólicas?

3. Discuta la proposición siguiente: Si se trata la muerte como la realización de las ganancias de capital del occiso, se eliminará una fuerza importante para la estabilidad del mercado de valores.

4. En Sanitary Farms Dairy, Inc., 25 T. C., 463, 467-468 (1955), el Tribunal Fiscal sostuvo que el propietario de una compañía de productos lácteos podía deducir de su impuesto al ingreso más de 15 000 dólares que había gastado en un safari africano que había realizado con su esposa. Durante el safari había tomado películas que más tarde utilizó extensamente en la publicidad de su negocio lechero. El tribunal sostuvo que "las pruebas revelan que una publicidad de igual valor a la implicada aquí no podría haberse ob-

tenido por la misma suma de dinero en ninguna forma más normal", y concluyó: "Nada de ese costo es gravable […] como gastos personales de viaje y diversión […] Es cierto que disfrutaron la cacería, pero el disfrute de nuestro trabajo no vuelve a ese trabajo una mera afición personal o el costo del viaje de cacería ingreso para el cazador". ¿Es este resultado económicamente correcto? ¿Debió el tribunal no haber permitido ninguna deducción? ¿Una deducción parcial?

5. ¿En qué circunstancias debieran ser deducibles del impuesto al ingreso los costos del servicio legal incidental del litigio?

6. En principio, ¿es un impuesto al ingreso necesariamente más eficiente que un impuesto específico? Véase Milton Friedman, *Teoría de los precios*, cap. 3 (1976).

7. ¿Sugiere la aversión al riesgo que un crédito fiscal para gastos médicos sería más deseable que una deducción fiscal destinada a generar, en total, el mismo ahorro tributario?

8. Es demasiado generoso que se permita a un comerciante deducir como gastos del negocio la cantidad total de cualquier impuesto a las ventas que remita al gobierno. Explique.

9. ¿Debiera permitirse que una empresa de servicio público dedujera sus contribuciones caritativas? ¿Que tratara tales contribuciones como una parte de su costo del servicio para los fines de la determinación de las tarifas?

10. Trace las gráficas XVIII.1 y XVIII.2 como impuestos planos antes que como impuestos específicos *ad valorem*.

11. Las cuestiones tributarias que la economía puede iluminar no surgen sólo en los casos de impuestos. Considere Hemenway *vs.* Peabody Coal Co. 159 F.3d, 255 (7th Cir., 1998). Las partes tenían un contrato de arrendamiento para la extracción de carbón en cuyos términos la compañía carbonífera aceptaba pagar a los arrendadores una regalía basada en "el precio medio de facturación del carbón extraído, removido y vendido" de las instalaciones arrendadas. Años más tarde, el Congreso estableció dos impuestos específicos para las compañías extractoras de carbón: uno para cubrir el costo de la restauración de la tierra explotada a su estado original o prístino y otro para cubrir el costo de los beneficios otorgados a los mineros que hubieran contraído la enfermedad del pulmón negro por trabajar en las minas. La compañía carbonífera sumó el monto del impuesto a las facturas que extendía a sus clientes, pero no trató esta cantidad como parte del "precio de facturación" al calcular la regalía debida a los arrendadores. ¿Era esto un incumplimiento del contrato? "En términos de la economía", dijo el tribunal de instrucción al sostener que sí era un incumplimiento,

> el establecimiento de un impuesto a los productores de carbón representa un costo de producción fijo. Cuando el productor de carbón opta por trasladar este costo —o una parte del costo— al consumidor, el costo representa un desplazamiento de la

curva de oferta de carbón. Si suponemos que la demanda de carbón permanece constante, el "precio de venta" del carbón en el mercado aumentará necesariamente.

Por supuesto, la demanda de carbón podría tener cierta elasticidad que impidiera a un productor de carbón pasar de inmediato el costo de un nuevo costo de producción *[sic]* al comprador, y el precio del carbón podría permanecer constante o aumentar sólo para cubrir una parte de los costos de producción incrementados a corto plazo. Pero, si esto ocurre, algunos productores de carbón no podrán permanecer en el mercado debido al incrementado costo marginal de la producción del carbón. Sin embargo, el análisis del mercado nos dice que, con la pérdida de los productores de carbón, o de la oferta, en el mercado, el precio aumentará finalmente para reflejar ese costo total medio de la producción —incluidos los impuestos establecidos— para alcanzar el equilibrio del mercado.

Sin embargo, en este caso la demanda de carbón en el mercado era tal que Peabody pudo trasladar todo el costo de los incrementados impuestos y honorarios a los compradores. En consecuencia, Peabody (y sin duda otros productores de carbón) provocaron un desplazamiento inmediato de la curva de oferta, lo que a su vez condujo a precios mayores del carbón en el mercado. En virtud de que Peabody se benefició de esta capacidad para aumentar el "precio de venta" del carbón a fin de cubrir los nuevos o incrementados costos de producción, los arrendatarios [*sic*: debiera decir arrendadores] que tienen derecho a regalías tienen igualmente derecho a un aumento de sus regalías ganadas correspondientes a este cambio del precio del mercado.

¿Es ese razonamiento correcto? ¿Apoya la conclusión de que la compañía carbonífera violó el contrato? ¿Qué le parece el argumento de la compañía carbonífera de que los arrendadores obtendrían una ganancia súbita si se calculara la regalía sobre la base del precio de factura con inclusión de los impuestos, porque ese método de cálculo significaría que, entre mayores fuesen los costos de la compañía carbonífera y menores sus ganancias, tendría que pagar más la compañía carbonífera por concepto de regalía?

12. En los Juegos de Rol Online Masivos para Múltiples Reproductores (MMORPG), tales como *Everquest, SecondLife* y *The Sims*, "miles de usuarios interactúan [en Internet] unos con otros bajo la apariencia de personajes de juegos de video, y lo hacen de modo persistente: muchas horas al día, todos los días, a lo largo del año". Edward Castronova, *Synthetic Worlds: The Business and Culture of Online Games*, 147 (2005). "Al involucrarse en el juego, los jugadores trabajan, fabrican o se ganan artículos, venden artículos a cambio de efectivo y emplean el efectivo para comprarse otros artículos." *Ibid.,* 649. Si desean tener más dinero ficticio de esta clase, lo pueden comprar en eBay. Por ejemplo, les podría costar un dólar en eBay comprar mil dólares de dinero ficticio para un MMORPG en particular. Suponga que es usted un jugador exitoso y, habiendo comprado 100 000 dólares en dinero ficticio por 100 dólares,

en el espacio de un año convierte esos 100 000 dólares en 250 000 en dinero ficticio, el cual podría vender en eBay por 250 dólares, cosechando así una ganancia de 150 dólares. Suponga que no lo vende, sino que continúa usando los 250 000 dólares en dinero ficticio dentro del juego. ¿Debería imponérsele un gravamen con base en el incremento de valor (en dólares reales), esto es, los 150 dólares?

XIX. LA TRANSMISIÓN DE LA RIQUEZA TRAS LA MUERTE

XIX.1. Impuestos a la sucesión (y las donaciones)

Para muchos estudiosos de la tributación, el gravar la riqueza al momento de la muerte ha parecido estar sustancialmente libre de los efectos de sustitución y regresividad de tantos otros impuestos. Y es cierto que la distribución medida de la riqueza es mucho más desigual que la del ingreso —y en el capítulo XVII vimos cuán desigual ha devenido la distribución del ingreso en este país—. "En 2010, el 1% de los hogares en la posición superior (la clase alta) poseía 35.4% de toda la riqueza atesorada en el sector privado, y los integrantes del siguiente 19% (el estrato de los administradores, profesionistas y propietarios de pequeños negocios) tenían en su poder 53.5%, lo que significa que tan sólo 20% de las personas eran dueñas de un destacado 89%, dejando no más de 11% de la riqueza en manos de 80% de la población ubicada en el polo inferior (los trabajadores asalariados y a sueldo). En términos de la riqueza financiera (el valor total neto menos el valor de la casa propia), 1% de los hogares situados en el polo superior tenían una proporción todavía mayor: 42.1%."[1] Pero el hecho de que la desigualdad de la riqueza se dé en mayor medida que la del ingreso es atribuible a que la seguridad social no se capitaliza al calcular la riqueza de los individuos; si lo fuera, la desigualdad de la riqueza sería mucho menor.[2] En efecto, si se midieran plenamente el ingreso y la riqueza, las dos distribuciones serían muy semejantes, porque todo el ingreso se capitalizaría (sobre todo como capital humano), mientras que todos los beneficios generados por el capital se tratarían como ingreso (por ejemplo, el ingreso de renta imputado que se deriva de vivir en una casa propia). Con todo, lo que es innegable es que la cantidad de los activos comercializables al momento de la muerte varía ampliamente entre las familias. Aun así, un impuesto sucesorio (esto es, aquel que se calcula con base en el valor neto de las propiedades del finado y se aplica al patrimonio en sí al momento de la transferencia, independientemente de quién o quiénes lo hereden y sin que éstos deban solventarlo), por oposición a un impuesto a la heren-

[1] G. William Domhoff, "Wealth, Income, and Power", www2.ucsc.edu/whorulesamerica/power/wealth.html.

[2] Jeffrey G. Williamson y Peter H. Lindert, "Long-Term Trends in American Wealth Inequality", en *Modeling the Distribution and Inter-generational Transmission of Wealth,* 9 (James D. Smith [comp.], 1980). Además, cada día aumenta el número de los trabajadores que tienen derechos de pensión privados creados (de sindicatos o compañías), lo que también debiera contarse como riqueza.

cia (esto es, aquel que deben pagar por su cuenta el beneficiario o los beneficiarios y cuyo monto depende del valor de los bienes que ya son propiedad de cada uno de éstos, y, en algunas jurisdicciones, de cuánto reciben de legado), es caprichoso porque omite la riqueza de las personas que se repartirán la herencia, todas las cuales podrían ser primos pobres. Y no es un generador potente de rendimientos gubernamentales. A la tasa impositiva actual, el impuesto sucesorio federal genera sólo alrededor de 1% de los rendimientos fiscales recaudados por el gobierno federal,[3] en tanto que a tasas más elevadas podría suministrar aún menos rendimientos a causa de los efectos de sustitución. Estos efectos a veces se pasan por alto, y se aboga por el impuesto sucesorio al definirlo como un impuesto perfecto: se dice que no puede distorsionar la asignación de recursos en virtud de que las personas no pueden escapar a la muerte. La analogía es con un impuesto de capitación, por ejemplo: un gravamen creado simplemente a partir del estatus de ser estadunidense, condición de la que a duras penas se podría escapar porque hallar un sustituto del estatus gravado requeriría optar por la expatriación, la cual resulta muy costosa para la mayoría de los estadunidenses.

Con todo, antes de llevar la analogía a un extremo, debemos considerar en primer término por qué algunos individuos mueren dejando riqueza en lugar de consumirla toda durante su vida. Es algo superficial la explicación obvia de que tales individuos no saben cuándo van a morir, de modo que deben conservar riqueza para el caso de que vivan más de lo que esperan; utilizando nuestra riqueza para comprar anualidades (un seguro de vida a la inversa, que paga al titular una suma fija o variable hasta el día de su muerte, sin ninguna acumulación), podemos asegurarnos de no dejar una herencia considerable al morir. Sin embargo, las anualidades no son la respuesta completa para las incertidumbres de la vida. La gente desea riqueza líquida para hacer frente a las demandas extraordinarias o para estar en condiciones de protegerse contra la inflación. En principio, una anualidad puede ajustarse a un índice contra la inflación y hacer préstamos para las necesidades de emergencia, pero esto requiere contratos complejos que quizás una persona que esté planeando su jubilación no entenderá o en los que no confiará plenamente.

Una segunda explicación para las herencias es la motivación del altruismo. Limitando el consumo a lo largo de su vida, una persona puede aumentar la riqueza de sus herederos. Los economistas no han llegado aún a un consenso sobre la importancia de la motivación de la herencia en relación con la motivación egoísta identificada en el párrafo anterior.[4] El análisis se complica por

[3] En la medida en que induce donaciones de los contribuyentes de ingresos altos a los de ingresos bajos, podría reducir la recaudación del Tesoro proveniente del impuesto al ingreso.

[4] Véase Michael D. Hurd, "Research on the Elderly: Economic Status, Retirement, and Consumption and Saving", 28 *J. Econ. Lit.,* 565, 617-629 (1990). Véanse también Wojciech Kopczuk y Joseph P. Lupton, *To Leave or Not to Leave: The Distribution of Bequest Motives,* Finance and Economics Discussion Series 2004-2033 (Board of Governors of Federal Reserve System, 2004);

el hecho de que anualizar una parte o aun el total de la riqueza del jubilado no es realmente incongruente con su deseo de dejar una herencia. Debido a la incertidumbre acerca de cuándo morirá su benefactor, los herederos tienen dificultades para repartir el consumo a lo largo de su ciclo vital. Esta incertidumbre puede eliminarse si el testador compra una anualidad que le dará cualquier nivel de ingreso que desee para su vida sin ningún residuo a su muerte, dejando el resto de su riqueza a sus herederos.

Si predomina la motivación egoísta, es posible que una tributación sucesoria onerosa no tenga grandes efectos de asignación. Si no existe la motivación de la herencia, cualquier riqueza acumulada es simplemente una forma de seguro contra las vicisitudes de la vida, y carece de todo valor cuando la vida termina. Pero, en la medida en que sea importante la motivación de la herencia —como lo es obviamente para algunos, especialmente los muy ricos (la disminución en la utilidad marginal del ingreso implica que los individuos muy acaudalados le adscribirán menos valor a su riqueza que los individuos menos acaudalados, y por ello los primeros serán más altruistas y los legados son una manifestación de altruismo intrafamiliar, la variedad más común de éste)—, la tributación sucesoria más pesada, al elevar el precio del consumo póstumo en relación con el consumo actual, reducirá el incentivo para ahorrar y aumentará el incentivo para consumir.

Resulta difícil saber cuán malo es esto, en virtud de que todos los impuestos tienen efectos de sustitución. Aunque una tributación sucesoria onerosa reducirá el incentivo para trabajar porque el beneficio de acumular dinero para un legado es más pequeño entre más pesado sea el gravamen fijado al legado, los efectos agregados de la tributación sucesoria de índole pesada sobre el trabajo, el ahorro y el consumo son probablemente ligeros. La mengua en el incentivo que tienen los testadores abrumados por ese gravamen para trabajar duro puede verse contrarrestada por un incremento en el incentivo de sus potenciales herederos para trabajar con ahínco; puesto que, dada la disminución de la utilidad marginal del ingreso, puede esperarse que las personas trabajen con más ahínco entre menores sean sus ingresos. Y, al ser abolido el impuesto sucesorio, los rendimientos que genera tendrían que ser remplazados por alguna otra clase de tributación, la cual podría causar más distorsiones que aquél.

La mayor objeción contra los impuestos a la herencia consiste en que pueden eludirse a un costo relativamente bajo,[5] con el resultado de que ni siquiera afecta en gran proporción a las personas que sí tienen un motivo para hacer un legado. Los testadores pueden transferir sus activos a sus hijos o a otros posi-

Laurence J. Kotlikoff y Fumio Hayashi, "Parental Altruism and *Inter Vivos* Transfers: Theory and Evidence", 105 *J. Pol. Econ.,* 1121 (1997).

[5] Los impuestos nominalmente onerosos sobre los patrimonios individuales en una sucesión, en este país y en otros, parecen no tener ningún efecto sobre la concentración de la riqueza. Véase Edward J. McCaffery, "The Uneasy Case for Wealth Transfer Taxation", 104 *Yale L. J.,* 283 (1994).

bles legatarios antes de su muerte, reservando el ingreso proveniente de sus activos para el decurso de su vida (el equivalente a una anualidad, que expira a la muerte del beneficiario y no deja nada para ser distribuido entre los legatarios), y minimizar la pérdida de control sobre los activos poniéndolos en manos de administradores fiduciarios bajo diversas restricciones. Es cierto que hay un impuesto para las donaciones, pero puede eludirse haciendo donativos moderados de propiedades cuyo valor se espera que se aprecie, o bien, lo que es una técnica relacionada, haciendo los donativos tan adelantados de la fecha en que el receptor los disfrutará (por ejemplo, a nuestros nietos en un fideicomiso hasta que lleguen a la mayoría de edad) que la operación natural del interés compuesto los hará crecer hasta sumas sustanciales, y el interés se gravará a una tasa menor que si el donante hubiera conservado el dinero hasta su muerte. El impuesto a la donación puede eludirse por completo haciendo una donación de capital humano, como ocurre cuando compramos una educación cara para nuestros hijos. Hay todavía más procedimientos para transferir riqueza que escapan a la tributación de las donaciones: un impuesto puede hacer que el padre se abstenga de dar a su hijo mayor un millón de dólares, pero no le impedirá instalar al hijo como vicepresidente ejecutivo de su compañía. Esta forma de la donación tiene un costo, al revés de lo que ocurre con una transferencia de efectivo simple. El nepotismo puede reducir la productividad de la compañía, creando una externalidad si la compañía no es enteramente de la propiedad del nepotista, caso en el que una parte de los costos de su nepotismo se desplazará hacia los otros propietarios.

El hecho de que los impuestos sucesorios y de donaciones no causen una asignación severamente errónea de los recursos no es realmente un punto a su favor. Un impuesto a los yates sería un mal impuesto (dejando de lado las consideraciones de la envidia), aunque pocas personas desearan yates, si el resultado fuese disuadir a esos pocos de comprarlos; porque entonces el impuesto crearía alguna distorsión en el consumo sin generar ninguna recaudación. Los impuestos a herencias y donaciones generan escasa recaudación porque los contribuyentes sustituyen la actividad gravada; la sustitución es un costo social.

Aun cuando es fácil evitar la tributación sucesoria, la decisión de saltársela está lejos de no conllevar costos. Un impuesto sucesorio es un gravamen a la riqueza, y la forma de evitar este último consiste básicamente en poseer menos riqueza. A ojos de los defensores del impuesto sucesorio, su principal mérito reside no en los rendimientos medianos que genera (si bien alcanzar 1% del total de los rendimientos tributarios a nivel federal no es poca cosa), sino en su efecto de reducir la magnitud de los legados, una repercusión juzgada digna de producirse porque recibir una cantidad muy considerable de dinero le granjea al receptor una ventaja injusta. (Sus defensores asimismo señalan que, al exceptuar los legados caritativos, el impuesto sucesorio alienta a las personas acaudaladas a desviar una parte de su riqueza, a menudo

una proporción muy grande, y destinarla no a su familia sino a causas de lo más nobles.) ¿Pero es la acción de heredar dinero en verdad más injusta que la de heredar neuronas y vigor físico y mental? Nos complace gravar el ingreso que genera nuestra dotación genética; ¿por qué no sentirnos complacidos con gravar el ingreso derivado de nuestra dotación pecuniaria?

Puede ser en apariencia necesaria la tributación sucesoria para impedir la creación, a lo largo del tiempo, de fortunas enormes que pudieran crear desasosiego político. En Inglaterra se produjo una concentración inenarrable de la riqueza, en parte por la exención de impuestos concedida a la herencia, pero la causa subyacente era la práctica de la primogenitura: de dejar virtualmente todo al mayor de los hijos. Cuando las herencias se dividen entre los herederos, aun las fortunas masivas tienden a disiparse en pocas generaciones. La primogenitura es común sólo cuando los activos principales de las herencias son indivisibles, de modo que dejar los activos a más de uno de los hijos conduciría a las deseconomías de la propiedad dividida (véase antes III.9). Esto se aplica a menudo a la tierra de cultivo, la forma principal de la riqueza durante la mayor parte de la historia inglesa, pero por supuesto no en la actualidad. Hoy en día, además, la creación de focos de poder privado funciona como una forma de poner un dique al omnímodo poder gubernamental y, por consiguiente, en efecto conllevaría la capacidad de auspiciar la libertad, mientras que en otros tiempos los focos de poder de los particulares engendraban inestabilidad política en virtud de que los gobiernos eran por lo general débiles. En cualquier caso, si su objetivo es el de limitar la expansión de las fortunas, los gravámenes fijados a la herencia (cuya imposición corre a cargo de los estados, no del gobierno federal) tienen más razón de ser que las tributaciones sucesorias. Los impuestos a la herencia gravan el acto de transmisión y recepción de la herencia, y por lo tanto se pueden calibrar de acuerdo con la riqueza poseída por quien o quienes recojan el legado.

La tributación sucesoria (o, mejor aún, el gravamen fijado a las herencias) puede tener cierto atractivo como el primer paso hacia un impuesto general a la riqueza, lo que podría defenderse con el argumento de que la mera suerte desempeña un gran papel en la mayoría de las grandes fortunas. Gravar la riqueza sólo a la hora de la muerte significa castigar al altruista; la persona que gasta toda su riqueza en sí misma elude el impuesto por completo. Si la persona rica gasta su riqueza antes y no después de que se acumule, elude también un impuesto general a la riqueza. La riqueza atañe tanto a la esfera social como al ámbito de los ahorros personales, de modo que gravarla es como recoger el maíz o el trigo antes de que madure el grano. No obstante, un impuesto al consumo, asunto abordado en el capítulo precedente, fomentaría la frugalidad, aunque al reducir el consumo del contribuyente posterior al gravamen podría inducirlo a ahorrar menos (más allá de que el gravamen en cuestión favorezca el ahorro por encima del consumo) a fin de contrarrestar el efecto del impuesto sobre su nivel de vida.

Algunos defensores de las tributaciones sucesorias se preocupan por la declinación en la movilidad social sufrida en los Estados Unidos. Tal como se apuntó en el capítulo XVII, las personas ricas se muestran cada vez más hábiles para garantizar que sus hijos y aun sus nietos permanezcan en el nicho de quienes reciben mayores ingresos, dejando un espacio más y más estrecho para que lo ocupen los hijos y nietos de los pobres. Con todo, es improbable que el problema del declive ocurrido en la movilidad social se resuelva tratando de poner un coto a los legados, ya que los individuos provistos de un caudal pueden transferir gran parte su riqueza durante su tiempo de vida. Para tener un impacto significativo sobre la transmisión de la riqueza al sucederse de las generaciones, un impuesto rígido a los legados tendría que verse acompañado de un gravamen rígido a las donaciones, y la palabra *donación* debería definirse de manera amplia para incluir el pago de la educación en escuelas privadas que reciba un niño. Y una tributación sucesoria y de donaciones en verdad rigurosa, aun si fuera viable, podría resultar indeseable debido al efecto desincentivador que tendría sobre el esfuerzo laboral sostenido por aquellas personas que se sienten motivadas, en un grado notable, a volverse ricas, animadas como están por el deseo de que sus hijos y nietos se encuentren en una situación más cómoda de como estarían en otras circunstancias.

Existe una presión para abolir la tributación sucesoria a nivel federal (el "impuesto a la muerte", lo llaman sus opositores), lo cual no deja de sorprender dado que el impacto de este gravamen recae casi por entero sobre los muy ricos, cuya participación en el ingreso y la riqueza ha ido engrosando a paso firme en los años recientes y ha alcanzado alturas sin precedentes en la historia moderna de los Estados Unidos. Por añadidura, la abolición del impuesto sucesorio acaso tendría efectos bastante limitados, puesto que los estados podrían responder incrementando sus propias tributaciones sucesorias, que tienen menos eficiencia que las correspondientes a escala federal en virtud de que es más fácil evadir al fisco mudándose de un estado a otro que exiliándose, y, de este modo, tales gravámenes afectan más las decisiones locales que la incidencia que tienen los impuestos federales. Por supuesto, la movilidad de los contribuyentes limitaría el nivel en que cualquier estado habría de incrementar su tributación sucesoria; aun así, podemos observar grandes diferencias en las tasas de impuesto al ingreso aplicadas por los estados, pese a la habilidad que muestran los contribuyentes para votar contra un gravamen sacando las uñas.

Más aún, en la actualidad la tributación sucesoria les proporciona a los herederos una base fortificada en lo respectivo al capital que les dan por legado. Esto es, en el caso de que más adelante vendan un activo de capital que se les legó, la base del costo para calcular el monto del impuesto a las ganancias de capital que deben será el valor del activo a la hora de la muerte del testador, y no el costo en que incurrió el testador para adquirirlo originalmente. Este descuento fiscal desaparecería si se aboliera la sucesión tributaria. De

esta manera, la eliminación de este gravamen sería contrarrestada en un grado insospechado por un aumento en el impuesto sobre las ganancias de capital que pagarían los herederos, e igualmente por un aumento en los gastos administrativos debido a que con frecuencia es difícil determinar el valor original de un activo que se adquirió muchos años antes. Una alternativa a la de abolir la tributación sucesoria sería abolir la base fortificada, que constituye una excepción arbitraria al gravamen fijado a las ganancias de capital.[6] A resultas de semejante eliminación, la tributación sucesoria generaría rendimientos más abundantes para el gobierno (el patrimonio en sucesión estaría sujeto a un gravamen de mayores dimensiones) y asimismo tendría un mayor alcance redistributivo: al parecer, en promedio, las ganancias de capital no realizadas en los patrimonios se incrementan según el tamaño de éstos, en una gama que va de 13% para los patrimonios por debajo de los dos millones de dólares a 55% para los patrimonios superiores a los 100 millones.[7]

En una curiosa mancuerna con el movimiento dirigido a abolir la tributación sucesoria a escala federal, muchos estados han abolido o modificado la "Regla contra las Perpetuidades" (una regla de carácter estatal antes que federal) con la finalidad de dar paso a la instauración del bien llamado "fideicomiso de dinastía" y, de esta forma, atraer al estado a las empresas de fideicomisos.[8] La Regla contra las Perpetuidades, según recordamos por el capítulo III, estipula que ningún interés futuro es válido a menos que se confiera en el lapso de 21 años después de acontecida la muerte de alguien que estaba vivo cuando dicho interés se creó. Así pues, la regla es una especie de denominación descaminada, en virtud de que no limita la vigencia de una condición establecida en un legado, sino que más bien limita la facultad del testador de asignar expresamente donaciones a sus descendientes remotos. En un fideicomiso dinástico, el que ha creado un fideicomiso cediendo una propiedad gira instrucciones al administrador fiduciario para que invierta el dinero en provecho de beneficiarios específicos, tales como los descendientes (no importa cuán remotos) de aquél durante todo el tiempo en que existan éstos. Dependiendo de la cantidad repartida por el administrador fiduciario entre los miembros de cada generación, el fideicomiso podría acumular a través del tiempo enormes activos mediante la sencilla operación del interés compuesto. Este instrumento es bastante reciente y, empero, ya se han colocado 100 000 millones de dólares en fideicomisos de dinastía. El motivo principal a que obedece el instrumento no es el de perpetuar el control sobre los activos del fideicomiso sino

[6] Véase Edward J. McCaffery, "A New Understanding of Tax", 103 *Mich. L. Rev.,* 807, 913-917 (2005).

[7] Véase Robert B. Avery, Daniel Grodzicki y Kevin B. Moore, *Estate* vs. *Capital Gains Taxation: An Evaluation of Prospective Policies for Taxing Wealth at the Time of Death* (Board of Governors of Federal Reserve System, 2013).

[8] Robert H. Sitkoff y Max M. Schanzenbach, "Jurisdictional Competition for Trust Funds: An Empirical Analysis of Perpetuities and Taxes", 115 *Yale L. J.,* 356 (2005).

el de evitar un impuesto federal "saltador de generaciones", fijado en 1986 y que rellenó una fisura legal.[9]

Los costos de transacción involucrados en los fideicomisos dinásticos pueden ser gigantescos, a causa del incremento habido en el número de beneficiarios a lo largo del tiempo. ¡Imaginemos cuántos descendientes podría tener una persona 100 años después de su muerte! El administrador de un fideicomiso dinástico incurriría en grandes costos al ponerse a rastrear a todos los beneficiarios y asignarle a cada uno un ingreso fiduciario. Y, si el número de beneficiarios aumentara más rápidamente que los activos del fideicomiso, a la larga la cantidad por asignar a cada beneficiario se volvería tan pequeña que los costos de administrar el fideicomiso devorarían la totalidad de sus activos.

XIX.2. Herederos interesados: el caso del heredero asesino

En lo correspondiente al derecho de la herencia, el asunto que reviste más trascendencia desde un punto de vista económico consiste en saber si un tribunal debiera rehusarse alguna vez a hacer cumplir los términos de un testamento, aparte de los casos obvios de fraude e incompetencia mental. Supongamos que A, mentalmente sano, hace un testamento en el que hereda la mayor parte de sus posesiones a su nieto B. B asesina a A. ¿Debiera permitirse que B haga cumplir el testamento?[10] La ley contesta negativamente. Puesto que los testamentos no contienen nunca una cláusula expresa que desherede al asesino del testador, pudiera pensarse que esta respuesta significa que los tribunales están sacrificando las intenciones de donación de los testadores en aras de la desistencia del homicidio, e indebidamente (y quizás innecesariamente) agravando el castigo que la legislatura prescribió para el homicidio.

Pero estas objeciones se disuelven cuando se utiliza un enfoque económico. Es improbable (para decir lo menos) que una persona que cree que hay cierto peligro de que alguien que designe en su testamento lo asesine, nombre a esa persona en su testamento, aunque podría hacerlo, pero añadiendo que el legado se invalida si la persona nombrada lo asesina; esto podría ser mejor aún que no designarla, porque haría más difícil que ésta se librara de ser detectada en el caso de asesinar al testador.

[9] Véase Max M. Schanzenbach y Robert H. Sitkoff, "Perpetuities or Taxes? Explaining the Rise of the Perpetual Trust", 27 *Cardozo L. Rev.,* 2465 (2006). Los autores explican que, "previamente a 1986, el impuesto sucesorio podía evitarse por vía de sucesivos intereses de vida, por ejemplo, dejando la propiedad de por vida al hijo de uno, luego al nieto de uno. Como la tenencia de por vida finaliza con la muerte y el impuesto sucesorio se aplica solamente a los intereses transferibles del finado, en el ejemplo anterior no habría gravamen alguno cuando, tras la muerte del hijo del cesionista o transferente, el interés del nieto del transferente se volviera posesorio". *Ibid.,* en 2476-2477 (citas omitidas).

[10] Riggs *vs.* Palmer, 115 N.E., 188 (N.Y., 1889).

La regla que prohíbe que el asesino del testador lo herede desempeña así la función ya conocida de proveer en un contrato o escritura de traspaso una cláusula implícita para resolver las contingencias remotas. Si esto es correcto, el efecto principal de la regla estriba en reducir simplemente el costo de redacción del testamento. La regla también debería aplicarse cuando existe un patrimonio pero no hay ningún testamento.[11] Pues la ley especifica al heredero o herederos, en ausencia de testamento, y de esta manera el heredero o los herederos tendrán el mismo incentivo para cometer un asesinato que el beneficiario de un testamento.

Supongamos ahora que el heredero mató a su testador por accidente; ¿debiera ser descalificado de la herencia de todos modos? La descalificación logrará disuadir a los individuos de cometer un asesinato disfrazado de accidente, pero podría generar una disuasión excesiva, sobre todo si la herencia es mayor que el pago óptimo de daños por el asesinato. Los daños óptimos pueden ser en efecto nulos si el accidente se debió a la conducta no negligente del heredero. Y los herederos podrían también eludir a sus testadores por miedo a que ocurriera un accidente y se les descalificara de obtener una herencia; ese comportamiento de desvío sería un daño para el testador. Pero si el accidente se debió a la negligencia del heredero, se fortalecerá el argumento a favor de aplicar la regla del heredero asesino. Si el homicidio negligente no es un delito, de modo que la única sanción sea el pago de daños, no habrá ninguna sanción por el asesinato negligente de nuestro testador, por lo menos si el homicida es su único heredero, ¡porque el juicio contra el homicida sería promovido por el propio homicida en su calidad de heredero! No obstante, enfoquemos el análisis del asesinato accidental del testador por un heredero no desde el punto de vista de la disuasión sino desde el de tratar de ejecutar las probables intenciones del testador. Si el accidente es resultado de la simple negligencia, el testador podría haber estado (de haber meditado en el asunto *ex ante*) dispuesto a perdonar al heredero y, así, permitirle heredar.

El heredero asesino es el ejemplo extremo de un fenómeno más general: el del heredero interesado. Incluso los niños que aman a sus padres, o a otros parientes que posiblemente tienen un patrimonio que heredar, podrían mantenerse lo más posible a distancia del patrimonio por legarse a su nombre en vez de a nombre de sus hermanos u otros legatarios potenciales, incluidas las fundaciones caritativas. La herencia es una oportunidad de cosechar beneficios; de ahí que auspicie la búsqueda de renta.[12] Podemos imaginar a unos herederos que entran en competencia por un legado incurriendo en costos que finalmente se engullen la mayor parte o incluso la totalidad del legado

[11] Esto es, cuando el inmueble se transmite por "sucesión intestada", lo que significa que un tribunal de sucesiones ordena que se distribuya el inmueble entre los herederos del testador según lo determine el estatuto estatal.

[12] Véase James M. Buchanan, "Rent Seeking, Noncompensated Transfers, and Laws of Succession", 26 *J. Law & Econ.*, 71 (1983).

esperado. Algunos de esos costos podrían ser costos de prestar servicios de considerable valor al testador potencial; otros podrían ser comentarios menospreciativos sobre un competidor, y otros más podrían ser los costos de simular que se ama al testador potencial, como en *El rey Lear*. De todos modos, es poco probable que el mercado de servicios provistos a los testadores potenciales funcione bien, ya que un testador siempre puede revisar su testamento. La posibilidad de hacerlo opera como una sanción contra un comportamiento borrascoso por parte de los herederos potenciales, si bien al mismo tiempo puede propiciar un comportamiento borrascoso por parte del testador potencial.

Se trata de un problema muy añejo. Pensemos en Labán y el fraude que comete contra su yerno Jacob en Génesis. Después de prometerle la mano de su hija menor, Raquel, en retribución de los siete años de servicio que le brindara, Labán engatusó a Jacob —cuyo trabajo había aumentado su caudal— para que se casara en cambio con su hija mayor, Lea. Jacob fue forzado a servir a Labán por otros siete años a fin de que se le permitiera desposar a Raquel. Un ejemplo moderno de semejante comportamiento fue objeto de confrontación en el caso Monarco *vs*. Lo Greco.[13] Cuando el demandante cumplió 18 años y quería abandonar el hogar y forjar su propia ruta en el mundo, su madre y su padrastro le prometieron que si se quedaba en casa y trabajaba en la granja familiar le legarían casi la totalidad de su propiedad (que era de tenencia conjunta) a él. Permaneció con ellos y trabajó con denuedo, recibiendo a cambio de ello solamente un cuarto por alojamiento y comida suficiente y algún dinero para gastos comunes. La granja se hizo próspera. Con todo, cuando el padrastro falleció 20 años más tarde, le legó la mitad de su participación en la granja a su propio nieto. El demandante se había sobado el lomo en la finca por una compensación nimia durante 20 años (sacrificando entre otras cosas las oportunidades de obtener una educación más allá de la educación media) con la expectativa de ser bien compensado a la muerte de su madre o su padrastro. La granja había rendido fruto, en parte gracias a que las ímprobas labores del demandante (su "capital en sudor") habían sido subcompensadas. ¿Contaba él con un remedio? Según lo exigía el Estatuto de Fraudes, la clase de contrato que había acordado con su madre y su padrastro debía haberse puesto por escrito, y no había sido así. No obstante, el tribunal determinó que el demandante podía obtener un reparo del nieto de cualquier forma porque éste se había enriquecido injustamente. El Estatuto de Fraudes obedecía al propósito de aligerar la carga que tenía el tribunal de decidir en materia de problemas estilo "ella dijo… él dijo" al hacer fallos sobre disputas contractuales. El trabajo del demandante que llenara las arcas del nieto se había llevado a cabo obviamente con la razonable expectativa de recibir compensación, y estas circunstancias suministraban una garantía adecuada de

[13] 220 P.2d, 737 (Cal., 1950).

veracidad que fungía de alternativa a un contrato por escrito o a una declaración de intenciones puesta sobre el papel.

XIX.3. La mano del muerto y la doctrina *Cy pres*

La imagen refleja del testador que omite una limitación de la conducta de sus herederos (por ejemplo, "pierdes tu herencia si me matas") es el testador que inserta en su testamento una cláusula que limita los usos a los que podrán destinarse los activos de su herencia. El problema de la "mano del muerto" que controla el uso de los recursos por los vivos surge cuando la muerte no se traduce en una transferencia clara a las personas vivas que les permita hacer con el dinero lo que quieran. Dado que una motivación para acumular un patrimonio sustancial podría ser la de proyectar influencia más allá de la muerte mediante el establecimiento de condiciones (quizá perpetuas) sobre el uso de los fondos de la herencia, una política de descartar las condiciones del testador tendría en algunos casos un efecto muy similar sobre el incentivo para acumular al de una tributación sucesoria onerosa. Pero, si se obedecieran siempre las condiciones de un testamento, especialmente las perpetuas, un resultado frecuente sería que los recursos controlados por ellas se emplearían ineficientemente. Las contingencias imprevistas que se materializaran después de la muerte del testador podrían requerir que los recursos se reacomodaran a fin de maximizar la eficiencia. Si no pudieran alterarse las condiciones del testamento, no habría manera de producir la reasignación.

El carácter del problema se ilustra con una controversia acerca de un parque donado a la ciudad de Macon, Georgia, por Augustus Bacon, un senador de los Estados Unidos, en representación de Georgia, que falleció a principios del decenio de 1900. El testamento del senador Bacon, elaborado durante la época de la legislación segregacionista que siguió al final de la Reconstrucción, estipulaba que el parque sería usado solamente por mujeres y niños blancos. En el decenio de 1960 se promovió un juicio contra la ciudad, alegando que el cumplimiento de la condición racial violaba la cláusula de protección igual de la Decimocuarta Enmienda. La Corte Suprema invalidó la condición,[14] tras de lo cual los legatarios residuales del testamento de Bacon promovieron un juicio para que se declarara: *1)* que la donación del parque era nula porque la ciudad ya no podría cumplir la condición racial del donativo, y *2)* que la propiedad era de ellos bajo la cláusula residual del testamento. Ganaron.[15]

A primera vista, parecería que este resultado vindica la política del cumplimiento de los deseos de los testadores, tal como los revelan las condiciones establecidas en los testamentos. Pero esto se vuelve dudoso en un exa-

[14] Evans *vs.* Newton, 382 U.S., 296 (1966).
[15] Evans *vs.* Abney, 396 U.S., 435 (1970).

men más detenido. Es posible que el senador Bacon haya insertado la condición racial a fin de inducir a la ciudad a administrar el parque. No hay ninguna indicación de que el propósito dominante de la donación haya sido la promoción de la discriminación racial antes que el de proveer un espacio recreativo para los habitantes de Macon. Bacon era un liberal según los estándares de su época, de modo que es plausible que, si lo hubieran consultado al respecto en la década de 1960, cuando las actitudes raciales eran diferentes, habría preferido que el parque siguiera siendo un parque, aunque abierto a los no blancos, en lugar de que sus herederos distantes se subdividieran la propiedad para darle un uso residencial o comercial. Esto es especialmente plausible porque la ciudad podría siempre recomprar la tierra a los herederos y continuar usándola como un parque… el que tendría que abrirse a los no blancos. En ese caso no se respetarían las intenciones discriminatorias del senador Bacon, y el único efecto de la invalidación de la donación caritativa sería el de conferir ganancias súbitas a los herederos, un resultado que presumiblemente no deseaba Bacon, puesto que no les legó a ellos la propiedad. Esto suscita la pregunta de por qué existía una cláusula residual. La respuesta es que el donativo a la ciudad podría quedar malogrado por otro motivo, tal como el de que la ciudad no deseaba contar con un parque.

El resultado en este caso es a un tiempo cuestionable en lo tocante a los principios fundamentales y contrario a la doctrina estándar de entonces (y de mucho antes, como estamos a punto de ver) y de ahora. El dilema que se plantea entre el cumplimiento de la intención del testador o la modificación de los términos del testamento de acuerdo con los cambios ocurridos en las condiciones desde su muerte es a menudo, como lo fue en el caso Macon, falso. Es probable que una política de rígida adhesión a la letra del instrumento de donación frustre los propósitos del donador y el uso eficiente de los recursos. En el propio caso de Macon, no estaba implicada ninguna cuestión seria de eficiencia porque, si la tierra fuera más valiosa como un parque que en un uso alternativo, la ciudad podría siempre comprarla a los herederos. Pero supongamos que el senador Bacon le hubiera donado a la ciudad un sanatorio para tuberculosos en lugar de un parque. En virtud de que declinó la incidencia de la tuberculosis, y los avances de la medicina volvieron obsoleto el método del sanatorio para el tratamiento de la tuberculosis, habría disminuido el valor de las instalaciones donadas en su uso planeado. Finalmente se habría puesto en claro que las instalaciones serían más valiosas en otro uso. Al revés de lo que ocurría en el caso Macon, no habría ningún obstáculo legal para continuar aplicando todas las condiciones de la donación. Pero sería muy probable que la aplicación fuese contraria a los propósitos del donante, quien con su donación trataba de contribuir a la cura de la enfermedad, no de perpetuar instalaciones inútiles.

Podría parecer que la discusión anterior equivale a negar la competencia del donador para equilibrar el valor de una donación perpetua contra el costo,

en términos de eficiencia, que tales donaciones imponen con frecuencia. Pero, dado que nadie puede prever el futuro, un donante racional sabe que sus intenciones podrían finalmente frustrarse por circunstancias imprevisibles, de modo que podría presumirse que acepta implícitamente una regla que permita la modificación de los términos de la herencia en caso de que un cambio imprevisto frustre su intención original. Esta presunción no es absoluta. Algunos donantes racionales, desconfiando de la capacidad judicial para alterar inteligentemente los términos de la herencia a la luz de condiciones modificadas, podrían preferir la asunción de los riesgos implicados en una adhesión rígida a los términos originales. ¿Debiera respetarse su deseo? Adviértase que ello volvería a los testamentos más rígidos que las constituciones, las que pueden enmendarse, aunque con dificultad. Sin embargo, las constituciones autorizan explícitamente la enmienda. La posibilidad de dar por nula la intención expresa de un testador invocando la doctrina *Cy pres* podría desalentar las donaciones caritativas.

Sin duda alguna corre un flujo paternalista en las decisiones que toman los tribunales de pasar por alto las condiciones formuladas en un testamento. No obstante, hay más bases para justificar el paternalismo en el caso de los testamentos que en el de los contratos. Los contratos se pueden modificar, pero una persona no puede modificar los términos de su testamento después de haber fallecido. Y, aun cuando la gente a menudo cambia sus testamentos, la transmisión de la riqueza al momento de la muerte es un acontecimiento que ocurre una sola vez en la vida, por lo que cada testador se halla falto de experiencia. Por añadidura, el testamento de una persona se elabora con frecuencia cuando ya está en una edad bastante avanzada, y, aunque no se encuentre senil (caso en el que el testamento no se llevaría a efecto), su capacidad de juicio puede no ser tan sólida como lo era cuando estaba más joven.

Procedamos ahora a abordar asuntos de doctrina: a la inversa de la decisión tomada en el caso Macon, donde el cumplimiento continuo de las condiciones de una donación a la caridad ya no es económicamente viable debido a la ilegalidad (en el ejemplo del parque) o a los costos de oportunidad (en el ejemplo del sanatorio), el tribunal, en lugar de declarar nula la donación y transferir la propiedad a los legatarios residuales o, en caso de no haber ninguno, al estado, autorizará a los administradores de la institución receptora a que destinen los activos a un propósito relacionado o aproximado (*Cy pres*: proveniente del francés *si près* u *aussi près,* y que significa literalmente "tan cerca de") ubicado dentro del alcance general de las intenciones del donante. El conde de Craven —quien tenía "en mente la triste y lamentable forma en que Dios Todopoderoso castigó al reino, pero en especial a las ciudades de Londres y Westminster, en los años de 1665 y 1666, con la peste y la gran mortandad, y la ingente necesidad que había de abrir un albergue para las víctimas de la peste, y de reservar un lugar para incinerar a los muertos"— había establecido un fideicomiso para la construcción de una casa para los

apestados y de un foso para los cadáveres. Cuando la peste negra ya no asolaba a los pobres residentes de St. Martin's in the Fields, el tribunal de justicia, que supervisaba los fideicomisos caritativos, permitió que el fideicomiso empleara algunos de sus activos para ayudar en el tratamiento de las personas aquejadas de otras enfermedades contagiosas. La doctrina *Cy pres* se podía haber empleado de manera similar en el caso Macon para justificar la medida de hacer caso omiso de la condición racial por su obsolescencia.

XIX.4. Los incentivos de las fundaciones caritativas

Aun cuando no ocurran contingencias imprevistas, las donaciones caritativas perpetuas plantean un problema económico que nos hace recordar la preocupación por separar la propiedad y el control en la corporación mercantil moderna. Una fundación caritativa que tenga un ingreso sustancial, a perpetuidad, de su dotación original es una institución que no compite en ningún mercado de productos ni en los mercados de capital, y que no tiene accionistas ni propietarios de alguna especie. Su consejo de administración fiduciaria es perpetuo y no rinde cuentas a nadie (excepto a sí mismo) por el desempeño de la empresa. (Aunque los procuradores generales estatales tienen autoridad legal sobre la administración de los fideicomisos caritativos, tal autoridad es en gran medida formal.) Y como, al mismo tiempo, ni los administradores del fideicomiso ni el personal directivo tienen la clase de derecho de propiedad sobre los activos o el ingreso de la fundación que generaría un fuerte incentivo para maximizar el valor, no operan aquí ni la zanahoria ni el garrote (*cf.* antes XV.2).

Los incentivos para la administración eficiente de los activos de la fundación podrían fortalecerse mediante una regla que obligara a las fundaciones caritativas a distribuir toda donación que recibieran, principal e interés, incluida la dotación original, dentro de un periodo de tiempo corto y especificado. La fundación no estaría obligada a terminar sus operaciones dentro del periodo; podría continuar indefinidamente. Pero tendría que recibir donaciones nuevas de tiempo en tiempo para evitar el agotamiento de sus fondos. Dado que es improbable que los donantes entreguen dinero a una empresa que carece de vigor, la necesidad de volver periódicamente al mercado en busca de donaciones caritativas daría a los fiduciarios y administradores de las fundaciones caritativas un incentivo del que ahora carecen para operar con energía. Las fundaciones —principalmente religiosas y educativas— que venden sus servicios en el mercado o dependen del apoyo caritativo continuo, de modo que están sujetas ya a algunas restricciones competitivas, podrían ser exentadas de la regla del agotamiento.

Las objeciones para la regla sugerida son que resulta innecesaria —los donantes están ya en libertad para limitar la duración de sus legados carita-

tivos— y que podría por lo tanto reducir los incentivos para hacer donaciones caritativas, dado que cualquier elemento que disminuya la utilidad de la donación al imponer condiciones sobre su durabilidad que no sean del agrado del testador reducirá su incentivo para realizar la donación. Un argumento en contrario es que muchas fundaciones perpetuas se establecieron en una época en que la fundación era una institución novedosa. Es posible que una persona que estuviera creando una fundación en esa época no pudiera prever el problema de la administración ineficiente e irresponsable que acaso afectara a una fundación perpetua a resultas del conjunto peculiar de restricciones (o mejor dicho de la ausencia de restricciones) bajo las cuales operan estas instituciones.

XIX.5. Legados para propósitos

Antes de pasar de los fideicomisos caritativos a los privados, debemos considerar por qué los fideicomisos caritativos han sido extraídos de una categoría más amplia, la de los legados para propósitos, y se les otorga un tratamiento favorable.[16] La creación de un fideicomiso caritativo por testamento es un legado dirigido a un propósito, tal como la educación o la investigación médica, por oposición a la creación de un fideicomiso privado, que es para el beneficio de individuos. ¿Por qué no podría permitirse que una persona cree por testamento un fideicomiso para propósitos que no sean caritativos, pero que sean enteramente legales, por ejemplo, un fideicomiso para los candidatos del Partido Demócrata, o para la fabricación de software de computadora, o para el consumo de caviar? Una posible respuesta es que un legado para propósitos requiere una inversión mayor de recursos judiciales y gubernamentales que un fideicomiso privado. Los legatarios residuales tienen un incentivo para argumentar que el fideicomiso no se está aplicando para los propósitos especificados y que, por ello, está condenado a fracasar, razón por la que ellos (los legatarios residuales) debieran recibir el principal, como sucede en el caso de los fideicomisos caritativos. Un tribunal deberá decidir si el fideicomiso está siendo aplicado a los propósitos especificados y, en caso de que no lo sea, si se le ha de ordenar que se aboque a tales propósitos o más bien ha de ser clausurado en favor de los legatarios residuales. La amenaza y, en ocasiones, la realidad de tal litigación se consideran dignas de sobrellevar si el propósito del fideicomiso es caritativo, en vista del carácter de bien público que posee la donación caritativa.

La respuesta general de la ley para los fideicomisos de propósitos no caritativos pero legales consiste en permitir que el administrador fiduciario ejecute los términos del fideicomiso, sin obligarlo a ello. Si decide no hacerlo, la

[16] Véase Adam J. Hirsch, "Bequests for Purposes: A Unified Theory", 56 *Wash. & Lee L. Rev.*, 33 (1999).

propiedad se revertirá a la herencia del testador. Este enfoque desalienta los legados para propósitos, ya que el testador no tiene ninguna seguridad de que el legado se llevará a cabo efectivamente. Con todo, el testador puede anhelar que el administrador fiduciario posea la capacidad y el tino para ejecutar esta orden, al percatarse de que él (el testador) no puede anticipar todas las circunstancias que podrían adquirir relevancia para hacer viable y deseable la implementación del legado. Así pues, puede tratarse de un enfoque provisto de solidez económica.

XIX.6. Condiciones de los fideicomisos privados: origen del fideicomiso contra el dispendio

Supongamos que un hombre deja dinero a su hijo en un fideicomiso, pero que el fideicomiso terminará si el hijo no se casa con una mujer de religión judía antes de cumplir los 25 años de edad. En tales casos, el enfoque judicial es el de rehusarse a hacer cumplir la condición si es poco razonable. En el ejemplo anterior, podría ser un factor diferencial que el hijo tuviera 18 o 24 años en el momento del legado, y que la población judía en el lugar donde vivía el hijo fuera de mayor o menor magnitud.

Este enfoque pareciera enteramente desprovisto de un fundamento económico, y no hay duda de que el criterio de lo razonable no resulta iluminador aquí. Pero consideremos las posibilidades de modificación que existirían si el donativo fuese *inter vivos* antes que testamentario. A medida que se aproximara la fecha de ejecución, el hijo podría presentarse ante su padre y persuadirlo de que una búsqueda diligente no había revelado a ninguna muchacha judía casadera que pudiera aceptarlo. Podría persuadirse al padre para que otorgara una extensión o relajara de otro modo la condición. Si el padre ha muerto, esta clase de "recontratación" es imposible, y no puede aplicarse la presunción de que la condición es razonable. Esto aconseja la aplicación del enfoque de la doctrina *Cy pres* en los casos de fideicomisos privados tanto como en los caritativos, excepto quizá cuando el testador rechazara expresamente un poder de modificación judicial.

La observación apenas formulada podría explicar también el hecho de que, aunque el propietario de una colección de arte fuera enteramente libre para destruirla durante su vida (omitiendo la doctrina de los derechos morales; antes, en XI.1), un tribunal podría considerar que no es razonable una condición de su testamento que ordena la destrucción de la colección.[17] Es posible que nadie conociera la condición y que la protesta general surgida cuando se descubrió hubiera persuadido al propietario de abrogarla, si hubiese estado vivo para hacerlo. Tales casos son raros, pero un caso común ocurre cuando

[17] Véase, por ejemplo, In re Scott's Will, 93 N. W., 109 (Minn., 1903).

un escritor ordena en su testamento que su albacea destruya sus manuscritos inéditos. Estas instrucciones se desobedecen de ordinario. Desde un punto de vista económico, ¿debieran desobedecerse? Es una pregunta difícil; pero una razón por la que podríamos pensar que se deberían obedecer es que, de lo contrario, los autores podrían destruir sus manuscritos prematuramente. En realidad, la objeción más fuerte contra estas intervenciones paternalistas es que, a largo plazo, a medida que los testadores se enteren de la política de los tribunales de negar el cumplimiento de las condiciones que los jueces consideran poco razonables, reducirán sus legados. Sólo quienes sienten muy fuertemente estas cuestiones insertan en sus testamentos tales condiciones que influyan en la elección del cónyuge. La negativa de los tribunales a hacer cumplir las condiciones no hará que tales personas cambien de opinión, sino sólo que cambien sus testamentos.[18]

La regla *Cy pres* como tal no se aplica en los casos de fideicomisos privados; otras tres doctrinas sí lo hacen. Ya hemos discutido la doctrina de las condiciones irrazonables. Otra doctrina aplicable es la regla que prohíbe las restricciones a la enajenación: si te vendo mi automóvil, no puedo sacarte una promesa exigible de que no lo revenderás, a menos que tal condición sea necesaria para proteger un interés de seguridad que retengo en la propiedad (es posible que no me hayas hecho el pago completo). Esta regla es aplicable a los legados. El mérito obvio de la regla, por lo menos tal como se aplica a los legados hechos en el momento de la muerte, es que impide el uso ineficiente de recursos que derivaría con frecuencia de ciertas contingencias imprevistas; su desventaja ya conocida es que debilita el incentivo para trabajar duro a fin de acumular riqueza que uno quiere legar en su testamento. Pero tiene también un mérito más sutil: reduce los costos de transacción porque las restricciones para la enajenación, como los derechos de la primera negativa, crean en efecto una propiedad dividida (véase antes III.10), aumentando así el número de las partes cuyo consentimiento debe obtenerse antes de que la propiedad pueda transferirse.

La regla que prohíbe las restricciones de la enajenación es incongruente con la regla que permite los convenios restrictivos perpetuos (véase antes III.8), ¿o no? Esta última regla se basa en un reconocimiento faltante en aquélla de que, si no puedes controlar el futuro, éste afectará tu conducta en el presente. La doctrina *Cy pres* representa una posición intermedia entre los enfoques ejemplificados por las dos reglas.

[18] ¿Debería cumplirse rigurosamente una instrucción asentada en el testamento del testador en el sentido de que su inmueble se le dé en obsequio a quienquiera que sea la consorte del nieto del testador, siempre y cuando sea judía? Por una parte, el hecho de que la condición se le aplique a un nieto en lugar de a un hijo significa que el testador está intentando ejercer control sobre el futuro más remoto, pero al mismo tiempo se le está otorgando a ese nieto un margen de elección más amplio porque no existe un límite de edad para él, tal como en el ejemplo con que comienza esta sección.

La Regla contra las Perpetuidades, discutida anteriormente, se puede relacionar con otras limitaciones de la "mano del muerto" que hemos explorado. No sólo es probable que los arreglos para el futuro distante conduzcan a un uso ineficiente de los recursos, generado por contingencias imprevistas, sino que ciertos intereses que no se materializan sino en el futuro distante pueden ser propiedad de personas aún no determinadas o ni siquiera nacidas, lo que vuelve difícil o imposible obtener el consentimiento para una modificación de los términos del legado. Véase antes III.11.

Un procedimiento similar a la restricción de la enajenación, pero que la ley trata mucho más amablemente, es el fideicomiso contra el dispendio. Creado de ordinario, aunque no siempre, por un testamento, un fideicomiso contra el dispendio impide que el beneficiario (o sus acreedores) disponga del cuerpo del fideicomiso (el principal). (El ingreso derivado del fideicomiso, cuando se paga al beneficiario, se convierte en su propiedad y puede por lo tanto ser embargado por acreedores.) A primera vista, este procedimiento parece una fórmula para el fraude y el derroche. Si el fideicomiso genera un ingreso sustancial, el beneficiario tendrá la apariencia de ser rico y podrá endeudarse grandemente, pues tendrá todos los incentivos para hacerlo, porque si se va a la quiebra no podrán sus acreedores tocar su riqueza a menos que tenga otros activos no exentos.

Pero estas críticas son superficiales. Quienquiera que preste guiado por la apariencia de riqueza (la trampa común del hombre astuto) es un tonto. El fideicomiso contra el dispendio no es un fraude contra los acreedores a menos que se oculte la característica alusiva al dispendio que posee. Si no se oculta, lejos de permitir que el beneficiario pida prestado, tal característica restringe su capacidad de obtener préstamos limitando la cantidad de dinero disponible para pagar a los acreedores si el beneficiario no lo hace (véase antes XV.5). Este procedimiento es paternalista, pero, al revés de lo que ocurre con el paternalismo estatal, está movido por el altruismo, de modo que es más probable que maximice la utilidad del receptor. Y, al revés de lo que ocurre con las restricciones impuestas a la enajenación, no aumenta los costos de transacción. Difiere poco de un fideicomiso para los hijos, motivado por la desconfianza en su capacidad para administrar competentemente sus finanzas. Los fideicomisos se basan en la desconfianza.

XIX.7. La porción de la viuda

Otra limitación del poder de un testador es la disposición, encontrada en las leyes de herencia de todos los estados, que le prohíbe desheredar por completo a su viuda.[19] Esta limitación tiene una justificación económica. Según

[19] Como tantas doctrinas legales específicas del género, la de la "porción de la viuda" ha sido privada del género: la porción de la viuda es ahora la porción obligatoria del cónyuge supérstite de cualquier género. ¿Debiera el cónyuge supérstite que trabaja a tiempo completo tener derecho a

sabemos, es probable que la riqueza del esposo en el momento de su muerte sea un producto, en parte, del trabajo de la esposa, aunque ella no haya tenido nunca un ingreso pecuniario (véase antes V.1). Sin una protección estatutaria contra la decisión de arrebatarles el derecho a heredar una parte del patrimonio de sus esposos, las mujeres podrían negociar con ellos una protección contractual (los contratos para hacer legados son legalmente exigibles). La disposición estatutaria minimiza los costos de transacción.

El análisis precedente apoya también la decisión de excluir del patrimonio del esposo, para propósitos referentes a la tributación sucesoria, de una parte —bajo la ley vigente el total— de la porción de la herencia correspondiente a la esposa. Una parte del dinero que ella hereda de su esposo representa su propia acumulación de ingresos (aunque de ordinario se trata de ganancias imputadas antes que pecuniarias). Es también probable que ella sea una persona anciana, y, si ella muriera poco tiempo después que su esposo, el patrimonio de éste estaría sujeto a la tributación sucesoria dos veces en un periodo de tiempo corto, lo que equivaldría a un incremento caprichoso en la tasa de la tributación sucesoria.

Por último, adviértase que, dado que la herencia de un cónyuge supérstite no es intergeneracional, no tiende a perpetuar las fortunas o a agravar la desigualdad de la riqueza.

Lecturas recomendadas

Hirsch, Adam J., "Spendthrift Trusts and Public Policy: Economic and Cognitive Perspectives", 73 *Wash. U. L. Q.,* 1 (1995).

———, y William K. S. Wang, "A Qualitative Theory of the Dead Hand", 68 *Ind. L. J.,* 1 (1992).

Kaplow, Louis, "A Framework for Assessing Estate and Gift Taxation", en *Rethinking Estate and Gift Taxation,* 164 (William G. Gale, James Hines y Joel Slemrod [comps.], 2001).

Kelly, Daniel B., "Toward Economic Analysis of the Uniform Probate Code", 45 *U. Mich. J. Law Reform,* 855 (2012).

Kopczuk, Wojciech, "Incentive Effects of Inheritances and Optimal Estate Taxation", 103 *Am. Econ. Rev. Papers & Proceedings,* 472 (2013).

Langbein, John H., "The Twentieth-Century Revolution in Family Wealth Transmission", 86 *Mich. L. Rev.,* 722 (1988).

McCaffery, Edward J., "The Uneasy Case for Wealth Transfer Taxation", 104 *Yale L. J.,* 283 (1994).

una parte de la herencia del cónyuge que no trabajó? Es difícil decirlo. Los activos del cónyuge que no trabajaba pueden haberse originado en parte por las aportaciones financieras que realizara el cónyuge trabajador; pero esas transferencias podrían haber sido una compensación por las contribuciones no monetarias que hiciera el cónyuge que no trabajaba a la producción y la utilidad del hogar.

Schanzenbach, Max M., y Robert H. Sitkoff, "Perpetuities or Taxes? Explaining the Rise of the Perpetual Trust", 27 *Cardozo L. Rev.,* 2465 (2006).
Stake, Jeffrey E., "Darwin, Donations, and the Illusion of Dead Hand Control", 64 *Tul. L. Rev.,* 705 (1990).
Tullock, Gordon, "Inheritance Justified", 14 *J. Law & Econ.,* 465 (1971).

Problemas

1. ¿Hasta qué punto se reducen o eliminan las dificultades discutidas en este capítulo cuando se sustituyen los impuestos a las sucesiones y donaciones por un impuesto acumulado a los receptores de donaciones y herencias?

2. Un hombre muere y en su testamento deja un millón de dólares en un fideicomiso para su gato preferido, con instrucciones de que el ingreso del fideicomiso se utilice para aumentar la comodidad y el placer del gato. Desde un punto de vista económico, ¿debiera respetarse el legado? *Cf.* Dareh Gregorian, "Screw the Pooch: Leona's Pup Loses $10M of Trust Fund", *N.Y. Post,* 16 de junio, 2008, http://nypost.com/2008/06/16/screw-the-pooch/.

3. Discuta los efectos económicos de la prohibición de que los individuos deshereden a sus herederos.

4. A le deja dinero a B, quien es insolvente y es perseguido por sus acreedores. ¿Debiera permitirse que B rechace el legado de A? Véase Adam J. Hirsch, "The Problem of the Insolvent Heir", 74 *Cornell L. Rev.*, 587 (1989).

5. Si usted es propietario de la *Mona Lisa* y pinta (indeleblemente) un mostacho en ella antes de regalarle la pintura a su sobrina, de tal modo que su valor de mercado disminuye grandemente, ¿debiera calcularse su impuesto a la donación al valor de mercado antiguo o nuevo? *Cf.* Citizens Bank & Trust Co. *vs.* Commissioner of Internal Revenue, 839 F.2d, 1249 (7th Cir., 1988).

6. ¿Cuál sería el efecto sobre el número y el monto de las donaciones si éstas se gravaran a la misma tasa que los legados?

7. Sabemos que los individuos tienen tasas de descuento (positivas). ¿Las tienen las naciones? Supongamos que el gobierno está considerando iniciar un proyecto que sólo rendirá beneficios en el año 2050. Al valuar los beneficios, ¿debiera el gobierno descontarlos a su valor presente?

8. ¿Qué tiene de errado el argumento siguiente: Cuando una persona muere, la utilidad marginal de su ingreso baja a cero? Así pues, la cláusula de un testamento que ordene la destrucción de una propiedad valiosa no debiera respetarse porque, aunque la propiedad no tendrá ningún valor para el testador, sí lo tendrá (por definición del hecho de ser una "propiedad valiosa") para otras personas, de modo que habrá una reducción del bienestar social si la propiedad se destruye.

9. ¿Debiera permitirse que una persona cree un fideicomiso contra el dispendio para sí misma, o sólo para alguien más?

10. Explique los efectos que sobre los legados, la utilidad de donantes y donatarios y el comportamiento de los donatarios, tiene el hecho de sustituir un impuesto al ingreso por un impuesto al consumo. Véase Barbara H. Fried, "Who Gets Utility from Bequests? The Distributive and Welfare Implications for a Consumption Tax", 51 *Stan. L. Rev.,* 641 (1999).

Sexta Parte

EL PROCESO LEGAL

XX. EL MERCADO, EL SISTEMA ACUSATORIO Y EL PROCESO LEGISLATIVO COMO MÉTODOS DE LA ASIGNACIÓN DE RECURSOS

XX.1. Comparación de la asignación legal y la del mercado; comparación de los sistemas jurídicos inquisitivo y acusatorio

Hemos visto que la cuestión final para la decisión de muchos juicios consiste en determinar cuál asignación de los recursos maximizaría la eficiencia. El mercado decide normalmente esta cuestión, pero el sistema jurídico es el que debe decidir en situaciones en las que los costos de una determinación del mercado excederían a los de una determinación jurídica. Los criterios de la decisión son a menudo los mismos, pero ¿qué diremos de los procesos de toma de decisiones? Aquí encontramos algunos paralelos sorprendentes, junto con algunas diferencias significativas.

Al igual que el mercado, el derecho (especialmente el derecho común) utiliza la igualdad de los precios y los costos de oportunidad para inducir a los individuos a maximizar la eficiencia. Cuando los daños compensatorios son el remedio para el incumplimiento de una obligación, el efecto de la responsabilidad no es el de obligar al cumplimiento de la ley, sino el de obligar al infractor a pagar un precio igual a los costos de oportunidad de la infracción. Si ese precio es menor que el valor obtenido por el infractor, se maximiza la eficiencia si éste lo comete, y el sistema jurídico lo alienta, en efecto, a cometerlo; si es mayor, la eficiencia requiere que el infractor no cometa el acto y de nuevo el remedio del pago de daños provee el incentivo correcto. Al igual que el mercado, el sistema jurídico confronta al individuo con los costos de su acto, pero deja que el individuo decida si incurrirá en tales costos. Aunque a veces se imponen sanciones —castigos— mayores (por ejemplo, véase antes en VII.2), normalmente se hace esto sólo cuando se requiere crear los incentivos económicos correctos. De manera similar, se emiten requerimientos en lugar de otorgar el pago compensatorio de daños cuando no queremos que la gente evada el mercado, puesto que los costos de transacciones del mercado son inferiores a los costos de transacciones legales, o cuando los daños son inconmensurables o, de otro modo, resultan inadecuados como remedio.

Tanto el proceso legal como el mercado se caracterizan por su impersonalidad; en términos económicos, por la subordinación de consideraciones distributivas. La mano invisible del mercado tiene su contraparte en el desinterés del juez, que se mantiene al margen. El método por el que se compensa

a los jueces y se forman las reglas de la ética judicial está diseñado para garantizar que el juez no habrá de tener ningún interés financiero o de otra índole en el desenlace del caso que se le presenta. La impersonalidad judicial se refuerza por la práctica de excluir de los juicios consideraciones que atañan no a la conducta de las partes sino a los merecimientos a que se ha hecho acreedora cada una (véase antes VIII.9). El individuo pobre no tiene derecho a exponer la pobreza como una razón por la que se le deba excusar de la responsabilidad, como tampoco puede el individuo rico apelar al sentimiento de solidaridad de clase que pueda animar al juez. En el escenario del mercado, los vendedores tienen un incentivo financiero para ignorar las consideraciones distributivas y de esta forma maximizar la eficiencia, tal como lo discutiremos en el capítulo XXVIII. Los vendedores pagan un precio por subordinar las consideraciones comerciales a las ideológicas, a menos que sus clientes tengan preferencias ideológicas que se plasmen en elecciones sobre el consumo, lo cual sucede en algunos casos en los que estos últimos, por ejemplo, se ven impelidos a apartarse de vendedores que discriminan por motivos raciales —o que quizá no discriminan por motivos raciales—, causan daño al ambiente o venden armamento a regímenes opresores. A diferencia de los vendedores (y de los consumidores, que a veces tienen que pagar más para ser condescendientes con vendedores políticamente correctos), los jueces pagan por lo común un precio escaso por dar rienda suelta a sus preferencias ideológicas, al menos en los casos cerrados, y ese precio bien puede ser contrarrestado por la utilidad que derivan de expresar esas preferencias.

En lo concerniente a su manejo administrativo, el proceso judicial tal como existe en los Estados Unidos, al igual que en otras naciones cuyo sistema jurídico se origina en el de Inglaterra (a saber, los otros países anglófonos), se apuntala en los individuos de la esfera privada que están motivados en lo económico por el interés propio, en vez de apoyarse en personas altruistas o en funcionarios. A través del abogado que contrate, la víctima de una conducta que podría ser ilegal: *1)* investiga las circunstancias que rodean al acto presuntamente ilegal, *2)* organiza la información obtenida por la investigación, *3)* decide si activará la maquinaria de la asignación legal, *4)* introduce información a esa maquinaria en una forma digerible, *5)* verifica la corrección de la información aportada por el demandado, *6)* si es necesario presiona en favor de cambios de las reglas de asignación seguidas por los tribunales, y *7)* está al pendiente de que se verifique la recaudación del fallo. De este modo, el Estado queda facultado para prescindir de una fuerza policiaca para proteger los derechos que otorga a los individuos el derecho común, los fiscales se liberan de la obligación de hacerlos cumplir y otro personal burocrático queda exento de la obligación de operar el sistema. Estos funcionarios tendrían una motivación menos alta que la de un demandante privado, porque su interés propio en lo económico se vería afectado sólo indirectamente por los resultados de casos particulares. El número de los empleados públicos implicados en la

protección de los derechos de acción privados es sorprendentemente pequeño si se considera el monto de la actividad regulada por las leyes que crean tales derechos, así como el número de los empleados públicos implicados en la operación del mercado es pequeño en relación con la actividad organizada por el mercado.

Este proceso acusatorio no es universal. Muchos países no anglófonos emplean el sistema "inquisitivo" (que no debe ser confundido con la Inquisición), donde el juez encabeza la actividad de reunión de pruebas y configuración de los problemas, mientras que los abogados desempeñan un papel subordinado: más que *kibitzers,** pero menos que principales en el proceso de litigación. Los recursos destinados al proceso acusatorio son reducidos, y éste es un ahorro social en la medida en que tales recursos contrarrestan, en lugar de operar para aumentar, la corrección de la determinación judicial (compárense los gastos en publicidad). Contra esto debe considerarse el hecho de que un sistema inquisitivo implica un sustancial traslado de la responsabilidad por la aplicación de la ley del sector privado al sector público: la relación de jueces a abogados es casi 10 veces mayor en un sistema jurídico inquisitivo que en los sistemas acusatorios de los Estados Unidos e Inglaterra.[1] Esto implica una pérdida de eficiencia si, como se supone generalmente, el sector privado es más eficiente que el sector público.

Las magistraturas en los sistemas jurídicos (inquisitivos) del derecho civil son (como lo son prácticamente todos los de origen no británico) magistraturas *de carrera*.[2] En contraste, la mayoría de los jueces estadunidenses y de otros lares angloestadunidenses se convierten en jueces solamente después de una carrera en alguna otra rama de la profesión legal, como la práctica privada o el seguimiento de una acción judicial, o bien en la enseñanza. Una magistratura de carrera es básicamente una rama del servicio civil. La designación y la promoción tienen lugar por mérito. La promoción es un rasgo crucial porque un individuo recién graduado de una escuela de derecho habrá de ocupar necesariamente el peldaño más bajo en la escala judicial al recibir su primer nombramiento y tendrá la expectativa de ascender a posiciones que impliquen responsabilidades más severas a medida que

* Expresión derivada del yidis que equivale al "espectador dado a hacer juicios impertinentes: alguien que observa y ofrece consejos o comentarios no solicitados, por ejemplo, en un juego de naipes". [N. del E.]

[1] Véase Richard A. Posner, *Law and Legal Theory in England and America*, 28 (1996) (cuadro 1.1). Sin embargo, sostengo allí que esa proporción es engañosa por lo que toca a Inglaterra, y que, si "juez" y "abogado" se definen funcionalmente antes que formalmente, la proporción inglesa es similar a la del sistema inquisitivo. La razón es que el abogado que ejerce en tribunales superiores desempeña en el sistema inglés un papel cuasijudicial.

[2] Véanse, por ejemplo, Nicholas Georgakopoulos, "Discretion in the Career and Recognition Judiciary", 7 *U. Chi. L. School Roundtable,* 205 (2000); J. Mark Ramseyer y Eric B. Rasmusen, "Judicial Independence in a Civil Law Regime: The Evidence from Japan", 13 *J. Law. Econ. & Organization,* 259 (1997).

adquiera más experiencia. La promoción es menos importante en un sistema de entrada lateral; el cargo de juez de un tribunal federal de distrito es un nombramiento judicial obtenido por antigüedad y la mayoría de los jueces de distrito no son, y no esperan ser, promovidos a un nivel más alto.

En un sistema en el que la promoción competitiva constituye la clave para obtener un incremento en el ingreso, mayor prestigio y la satisfacción laboral, como ocurre en las judicaturas del derecho civil,[3] es difícil atraer solicitantes altamente calificados a menos que haya objetividad en los criterios para la promoción. Hay obstáculos para valorar la producción de quien se desempeña en el gobierno como burócrata, incluido un juez. Las burocracias (incluida la judicial) buscan por lo tanto generar estándares de promoción que tengan credibilidad y minimizar los costos de agencia estableciendo reglas detalladas para su acatamiento por parte de los burócratas, y es más fácil determinar el apego a una regla que el grado en que el burócrata se muestra como un sujeto creativo, innovador o imaginativo. Por consiguiente, las magistraturas de carrera se hallan en sistemas jurídicos que se apoyan considerablemente en códigos en vez de en la elaboración de las leyes por parte de jueces capaces de aplicar un criterio amplio, tal como sucede en los sistemas del derecho común.

Además, en una magistratura de carrera, al igual que en cualquier otra rama del derecho civil, la promoción depende de que uno demuestre poseer habilidad para desempeñarse de acuerdo con las expectativas de sus superiores, quienes probablemente no saldrán beneficiados en sus propias carreras de contar con subordinados audaces, propensos a aventurarse en rutas experimentales. En una empresa mercantil, un jefe de departamento que cuente con empleados innovadores y caracterizados por acometer con ahínco sus tareas y, así, contribuir al aumento de los rendimientos y ganancias verá cómo su crédito se eleva por haber seleccionado y alentado a esos subordinados. Esa forma de evaluar los resultados reales obtenidos en la cuenta final no está disponible para las magistraturas. De esta manera, es de esperar que los frutos rendidos por una magistratura de carrera apenas si presenten altibajos que sean de una calidad profesional característicamente uniforme, pero que a la vez estén faltos de creatividad. Así, no debería sorprendernos que en las naciones donde existen magistraturas de carrera se les dé a los profesores de derecho no sólo el trato de comentadores de la ley, como en los sistemas angloestadunidenses, sino también el correspondiente a quienes son fuentes de la ley. Los jueces están indispuestos a desempeñar ese papel.

En los sistemas de derecho civil tal como los que hemos descrito, los criterios de desempeño que hacen hincapié en la creatividad e independencia intelectual, en la sabiduría política o incluso en la habilidad para captar a

[3] Véase, por ejemplo, Martin R. Schneider, "Judicial Career Incentives and Court Performance: An Empirical Study of the German Labour Courts of Appeal", 20 *Euro. J. Law & Econ.,* 127 (2005).

profundidad y de forma pragmática el modo de operar de la "ley en acción" estarían fuera de lugar. Ésa es una razón de que el movimiento de derecho y economía haya tenido un impacto mucho más reducido en la práctica judicial de la Europa continental que en los Estados Unidos. Otra es que una magistratura que no es de carrera se halla mucho mejor equipada para emitir una sentencia sólida desde el ángulo económico.[4] Los jueces de entrada lateral tienen más experiencia con la sociedad comercial que los burócratas y se basan más en la información proveniente de la esfera situada más allá de la gubernamental (sobre todo la de los abogados, que reflejan las preocupaciones de sus clientes). Debido a estas características, los jueces angloestadunidenses tienden a mostrarse más solícitos y a estar más enterados respecto a los derechos de propiedad y otras características de una sociedad comercial que los jueces de carrera, y en particular es más probable que deriven principios jurídicos de los usos y costumbres comerciales.[5]

La función que realiza un sistema jurídico al asignar recursos haciendo cumplir las reglas de conducta, en lugar de simplemente resolver disputas *ad hoc*, sugiere que hay una justificación económica para que el gobierno sufrague una parte de los costos del sistema (los salarios de los jueces, el costo de la construcción y el mantenimiento de los tribunales, etc.). Si los tribunales solamente se dedicaran a zanjar disputas que vienen a planteárseles en busca de resolución, el criterio de eficiencia dictaría que es a los litigantes (a excepción de los demandados por un delito que son indigentes, según se explica después en XXIV.3) a quienes les correspondería cargar con el costo total del sistema. No obstante, en virtud de que el sistema también establece reglas que moldean la conducta futura, los beneficios sociales de una demanda pueden sobrepasar los beneficios privados de los litigantes, y, de este modo, la cantidad correspondiente a la litigación podría de hecho ser demasiado pequeña si los litigantes tuvieran que soportar los costos totales.[6]

La otra cara de la moneda es que, entre menos peso tengan las decisiones como precedentes —es decir, como reglas de conducta—, serán menores los beneficios externos que la litigación confiera y, por ende, será más débil el argumento a favor del subsidio público de ésta. Por consiguiente, sería menos de esperar que hubiera litigaciones en los sistemas legales inquisitivos porque éstos se apoyan más en los códigos y menos en el precedente que los sistemas acusatorios. Surge un contraste sugerente entre la noción angloestadu-

[4] Véanse Paul G. Mahoney, "The Common Law and Economic Growth: Hayek Might Be Right", 30 *J. Legal Stud.,* 503 (2001); Simeon Djankov *et al.*, "Courts", 118 *Q. J. Econ*., 453 (2003); Edward Glaeser y Andrei Shleifer, "Legal Origins", 117 *Q. J. Econ*., 1193 (2002); Michael L. Smith, "Deterrence and Origin of Legal System: Evidence from 1950-1999", 7 *Am. Law & Econ. Rev.,* 350 (2005).

[5] Todd J. Zywicki, "The Rise and Fall of Efficiency in the Common Law: A Supply-Side Analysis", 97 *Nw. U. L. Rev.,* 1551 (2003).

[6] Sobre la demanda y oferta de reglas legales, véase después XXI.2.

nidense de *stare decisis* y la de *jurisprudence constante,* que es propia de las magistraturas del derecho civil, tales como las de la Europa continental, y que no le da a una decisión el peso del precedente hasta que esa decisión se ha mantenido en una serie de casos subsecuentes. El juez en el segundo caso no está ni siquiera presuntamente obligado a apegarse a la decisión del primer caso, pero si está de acuerdo con ella y por lo tanto la mantiene, y si proceden similarmente los jueces de varios casos subsecuentes, entonces la decisión (o la secuencia de decisiones) posee fuerza como precedente, lo que significa que el solo hecho de que se haya adoptado la decisión tal como se tomó constreñirá a los jueces a mantenerla, al menos presuntamente.

Esta diferencia puede asimismo guardar relación con el mayor grado de uniformidad que exhiben los jueces en un sistema continental merced a la estructura más rígidamente jerárquica que distingue a éste. Los jueces que se apartan de los "lineamientos de la empresa" no son promovidos. A resultas de ello, la fuerza disciplinadora consistente en exigir a los jueces que se adhieran a las decisiones previas es menos importante.

Por añadidura, hay un trueque de compensaciones y sacrificios entre, por un lado, la certidumbre jurídica, que es auspiciada por el sistema angloestadunidense en virtud de que en ella hay menos tiempo de espera antes de que una decisión se convierta en el precedente, y, por otro, la evitación de los errores judiciales, que es auspiciada por la práctica continental de demorar el reconocimiento de una decisión como precedente hasta que haya suscitado por lo menos un consenso judicial preliminar. La diferencia en la manera como los dos sistemas sopesan las ventajas y desventajas de este intercambio es congruente con el hecho de que el sistema continental cuente con fuentes de certidumbre jurídica (códigos que abarcan más elementos y una jerarquía judicial más férrea) de las que el sistema angloestadunidense carece.[7]

Un contraste aún más agudo que el habido entre los sistemas jurídicos acusatorio e inquisitivo y judicial es el que se da entre los jueces (árbitros) públicos y privados. Se recurre a los árbitros para dirimir muchas disputas contractuales, incluidas la mayoría de las disputas surgidas a partir de los contratos en las negociaciones colectivas. Ya que el Estado no paga ninguna parte del gasto por el arbitraje, no debería causar azoro que la mayor parte de los árbitros no pongan por escrito sus opiniones, cuyo valor va a parar a manos principalmente de personas ajenas a las partes envueltas en el arbitraje (individuos que no han de contribuir a estos gastos). Algunos árbitros anotan sus opiniones, no obstante, presumiblemente a modo de publicidad, de lo que se colige que se trata sobre todo de personas que se especializan en el arbitraje. Tal especialización es más común en el arbitraje laboral que en el comercial, y en realidad es más probable que los árbitros laborales pongan por escrito sus opiniones que los árbitros comerciales.

[7] Véase Barbara Luppi y Francesco Parisi, "Rules *versus* Standards", en *Production of Legal Rules,* 43 (Francesco Parisi [comp.], 2011).

Otra diferencia entre los árbitros y los jueces radica en que un árbitro que se hace célebre por favorecer a una parte o a otra en una clase de casos, por ejemplo, los casos de terminación de empleo o las disputas entre inversionistas y corredores o entre administradores y sindicalistas, será considerado inaceptable por una de las partes en una disputa de esa naturaleza, y, en consecuencia, la demanda de sus servicios habrá de menguar. Podemos esperar, por consiguiente, que entre los árbitros se observe una tendencia a "dividir en dos mitades la diferencia" al emitir sus fallos, esto es, a concederle a cada parte una victoria parcial (y por ello una derrota igualmente parcial).[8] Pues entonces les será más difícil a las partes ubicadas a cada lado de la clase de demandas en cuestión discernir un patrón de favoritismo. Más aún, la pauta usada les será atractiva a los contendientes que son adversos al riesgo (dado que eliminará a la vez el riesgo en sentido ascendente y descendente que entrañe el proceso de resolución de la disputa) y, por lo tanto, contribuirá a diferenciar entre el arbitraje y la adjudicación. Esto encierra importancia porque el subsidio otorgado por el gobierno a los tribunales coloca a los árbitros en una posición de desventaja respecto a los costos. Una manera de superar esta desventaja consiste en ofrecer un servicio distintivo, y la toma de decisiones fundada en la práctica de dividir en dos mitades la diferencia es justamente un servicio de esa índole. Adviértase, empero, que las partes enfrascadas en la litigación en ocasiones disponen de un margen para elegir a los jueces. Un demandante puede tener posibilidades de seleccionar en qué tribunal se entablará la demanda, y un demandado puede imponer su propia elección solicitando una moción para transferir el caso a un tribunal más conveniente.

Volviendo a las diferencias entre el derecho y el mercado como métodos de asignación de los recursos, nótese que el mercado es un mecanismo más eficiente para valuar usos rivales. En un mercado, los individuos tienen que respaldar sus aseveraciones de valor con dinero, o con algún sacrificio equivalente de oportunidades alternativas. La disposición a pagar imparte mayor credibilidad a una reclamación de valor superior que la energía forense. Las dificultades de la determinación judicial de las preferencias o los valores relativos podrían explicar no sólo la renuencia de los tribunales (señalada a lo largo de la Segunda Parte del presente libro) a hacer juicios de asignación cuando los costos son bajos, sino también la tendencia judicial a evitar hacer juicios de asignación de envergadura, y punto. Recuérdese que al decidir si un victimario es negligente, los tribunales consideran si el cuidado que mostró fue razonable, pero no consideran si el accidente podría haberse evitado a un costo menor al costo esperado del accidente realizando otra actividad (véase antes VI.5).

[8] Véanse, por ejemplo, Armendariz *vs.* Foundation for Health Psychcare Services, Inc., 6 P.3d, 669, 693 (Cal., 2000); Donald Wittman, "Lay Juries, Professional Arbitrators, and the Arbitrator Selection Hypothesis", 5 *Am. Law & Econ. Rev.,* 61, 81 (2003).

Otra consecuencia de la incapacidad del derecho para medir correctamente las preferencias es una tendencia a suprimir las variaciones del valor. Muchas personas asignan a sus casas un valor que supera su precio de mercado. Pero un estándar de valor subjetivo en los casos de dominio eminente, aunque sea el estándar correcto en cuanto al principio económico, sería virtualmente imposible de administrar en virtud de la dificultad existente para probar (excepto por la prueba de la negativa del propietario a aceptar una oferta de buena fe justo por debajo de lo que él afirma que es su valuación subjetiva) que la casa valía más para el propietario que el precio de mercado (véase antes III.6).

Se ha sostenido que una diferencia aún más fundamental entre las transacciones legales y las del mercado es que quien transfiere no se ve compensado en el primer caso y sí en el segundo.[9] Si A le compra a B un automóvil, debe pagarle por supuesto a B por ello. Pero si A destruye el automóvil de B en un accidente en el que ninguna de las partes tiene culpa, en efecto se habrá permitido que A use el automóvil de B como un insumo de la conducción, sin tener que pagarle a B. Sin embargo, el contraste está exagerado. También en el mercado hay muchos perdedores. Si fabrico coches de caballos y el mercado de mi producto se derrumba cuando se inventa el automóvil, no seré compensado por la pérdida de mi negocio. Esa pérdida es eficiente, pero sólo porque las ganancias superan las pérdidas cuando se suman en toda la sociedad, no porque obtenga yo una parte de las ganancias suficientemente grande para cubrir mis pérdidas. En efecto, desestimar la competencia como un cuasidelito, como ocurre en este ejemplo, es un ejemplo claro de la subordinación de las consideraciones distributivas a las de la eficiencia en el derecho común. Del mismo modo, si existen buenas razones para requerir una prueba de negligencia en el caso de un accidente automovilístico, la pérdida de B es compensada por las ganancias que obtiene la sociedad en su conjunto.

Si esto es "injusto", no lo es más que en la competencia de mercado. Y los individuos *se ven* a menudo compensados —por adelantado— por las pérdidas por las que la ley se rehúsa a compensarlos. Supongamos que un sistema de negligencia es, en efecto, un sistema más barato para controlar los accidentes automovilísticos que un sistema de responsabilidad estricta. Entonces, la suma de la responsabilidad de B y de las primas de los seguros de accidentes debiera ser menor con un sistema de negligencia. Si resulta lesionado en un accidente en el que ninguna de las partes tiene culpa, B será compensado con cualquiera de los sistemas: por su compañía de seguros de accidentes (de primera parte), con el régimen de negligencia, y por la compañía de seguros de responsabilidad del victimario, con la responsabilidad estricta, en la que él tiene en efecto una reclamación legal. Pero, por hipótesis, será compensado a

[9] Jules L. Coleman, "Efficiency, Utility, and Wealth Maximization", 8 *Hofstra L. Rev.*, 509, 541-542 (1980).

un costo menor para él con un sistema de negligencia en el que debe buscar a su aseguradora para obtener la compensación.

XX.2. Comparación de la asignación judicial y la legislativa

Las reglas elaboradas por los jueces tienen más probabilidad de promover la eficiencia que las que realizan las legislaturas, más allá de las que producen éstas para codificar los principios del derecho común (por ejemplo, una gran parte del derecho penal y muchas disposiciones del Código Comercial Uniforme, así como otras leyes con carácter de uniformidad). ¿Cómo se explica esta diferencia importante en el carácter de la ley judicial y la legislativa?[10]

Una explicación reside en las diferencias existentes entre los procedimientos por los cuales se formulan reglas de derecho por los jueces o por los legisladores. Como ya se explicó, es improbable que un juez, especialmente de un tribunal de apelación, que es donde se hacen la mayoría de las reglas elaboradas por los jueces, decida un caso basado en la consideración de cuál de las partes es la persona "mejor". El juez conoce a las partes menos aún que el juez de instrucción; y, como se mencionó antes, las consideraciones referentes a los merecimientos relativos de las partes (riqueza, pobreza, buena crianza, etc.) se suprimen hasta donde ello sea posible. Además, un juicio basado en tales consideraciones sería difícil de racionalizar en una opinión judicial. Por último, los métodos de la compensación judicial y las reglas que gobiernan los conflictos de interés excluyen una elección entre las actividades rivales basada en el interés propio estrictamente económico del juez. El juez está obligado a ver a las partes como representantes de actividades: ser propietario de la tierra, cultivar tulipanes, caminar por las vías del ferrocarril, conducir automóviles. En estas circunstancias, es natural que se deba preguntar cuál de las actividades rivales es más valiosa en el sentido económico.

Las reglas de la adjudicación dificultan que los grupos de interés sean oídos siquiera. El concepto de la "capacidad para presentar demandas" hace que el derecho a demandar esté limitado en una corte federal y en la mayoría de los tribunales estatales a una persona u organización que pueda demostrar que obtendrá una ganancia específica y tangible del juicio si lo gana. Abolir el requisito aumentaría en forma desproporcionada la cantidad de litigios y también reduciría la calidad de las adjudicaciones, porque las cortes sólo podrían especular acerca de las consecuencias de una decisión para una u otra parte si, por ejemplo, el demandante estuviera cuestionando la validez o la interpretación de un estatuto recientemente promulgado. Por añadidura, el requisito de la capacidad para presentar demandas es el método por el cual se

[10] Sobre las razones que dan cuenta de la tendencia del derecho común a la eficiencia, véase también antes VIII.1 y después XX.5, XX.6 y XXII.6.

asignan los derechos de propiedad a los reclamos legales. Supongamos que A está acosando a B, en Nueva York, debido a que B adopta una posición "políticamente incorrecta" sobre cuestiones de raza y género. C, que vive en California, se entera del acoso, está indignado y demanda a A, en busca de un fallo. No obstante, B desconfía de la habilidad de C para procesar la demanda hasta llegar a una conclusión satisfactoria para él (B). Tal vez C sea una empresa legal de "interés público" adscrita a una tendencia conservadora que tenga un mayor interés en dramatizar el asunto de la corrección política que en obtener un fallo que permita salvar a la empresa de B. El requisito de la capacidad para presentar demandas le impide a C levantar una, y, por este efecto, coloca el derecho exclusivo a demandar en las manos de la persona a quien ese derecho le resulta más valioso.

El proceso legislativo muestra un contraste marcado con el judicial. No hay ninguna regla en contra de que se tomen en cuenta los merecimientos de las personas afectadas por la legislación propuesta. No se emplea el sistema acusatorio, con su comparación de actividades concretas que se interfieren, lo que asegura que las cuestiones de los costos relativos están siempre cerca de la superficie de la controversia. Sobre todo, las herramientas legislativas para la redistribución de la riqueza son mucho más flexibles y poderosas que las judiciales. De ordinario, la única forma como un tribunal de derecho común puede redistribuir la riqueza es por medio de (lo que es en efecto) un impuesto específico a la actividad implicada en el juicio. No es fácil redistribuir la riqueza por este medio.[11] Y es difícil imaginar tribunales del derecho común que creen una ley del derecho común con la finalidad de redistribuir la riqueza. Resultaría intrincado para los tribunales obtener la información necesaria para saber en qué punto fijar el salario, y, más importante aún, se verían notablemente presionados para justificar su acción al crear un salario mínimo excepto por razones distributivas que no suscitan un consenso; por consiguiente, la acción de los tribunales sería percibida como descaradamente política.

Ésa podría haber sido la razón de que el crecimiento del moderno Estado de bienestar esperara al advenimiento de la tributación del ingreso (por supuesto, la causalidad podría operar en la dirección contraria). Con un impuesto al ingreso establecido, la redistribución se alcanza más eficientemente elevando las tasas del impuesto al ingreso que adoptando normas jurídicas ineficientes. Aunque tanto los impuestos específicos (o las normas que funcionan como impuestos específicos) como los impuestos al ingreso distorsionan la asignación eficiente de recursos, como sabemos —y supongamos que lo distorsionan en la misma medida—, que el costo del peso muerto de una norma ineficiente puede superar el costo del peso muerto del pequeño aumento

[11] Véase antes XVII.7; véase también Raaj Kumar Sah, "How Much Redistribution Is Possible through Commodity Taxes", 20 *J. Pub. Econ.,* 89 (1983).

de las tasas del impuesto al ingreso que sería necesario para lograr la misma cantidad de redistribución que la norma jurídica.[12]

Por lo tanto, si los tribunales del derecho común no tienen herramientas eficaces para la redistribución de la riqueza, será en beneficio de todos los grupos de interés que los tribunales, cuando estén aplicando los principios del derecho común, se concentren en el esfuerzo de aumentar la riqueza agregada de la sociedad logrando que los principios y los resultados del caso sean eficientes.

Todo esto no es para sugerir que la sociedad podría gobernarse eficientemente sólo por el derecho elaborado por los jueces. El requerimiento de tener un fundamento jurídico limita las ocasiones en las que un tribunal puede elaborar o modificar una norma. No puede actuar cuando se hace evidente por primera vez la necesidad de una norma nueva o una modificación de una antigua; debe esperar hasta que una disputa legal concreta llegue al tribunal para resolución. Esto no servirá cuando la nueva ley debe elaborarse deprisa, o cuando se desea un cambio *grande* en la ley. Ésta es una razón de suponer que un estatuto (ley) se aplica en retrospectiva, mientras que una nueva doctrina elaborada por los jueces se aplica retrospectivamente. Otra razón es que un cambio de la ley elaborada por los jueces tiene mayores probabilidades de ser previsto, porque los jueces articulan y tienden a adherirse a principios establecidos, antes que a responder simplemente a los cambios de la opinión pública. Los jueces que no están sujetos al control democrático mediante el proceso electoral podrían considerarse agentes infieles de la ciudadanía, incluso tiranos, a menos que demuestren que están operando dentro del campo de la autoridad delegada. Por ambos factores (el alcance y la anticipación), un cambio en la ley elaborado por los jueces deja sin satisfacer menos expectativas que un cambio legislativo, aunque, por supuesto, esto sólo es cierto en la media de los casos. Algunos estatutos se aplican retroactivamente[13] y algunos cambios judiciales a la ley sólo prospectivamente por el efecto que la aplicación retrospectiva habría tenido en las expectativas acordadas. Adviértase que los tribunales serían menos "activistas" si siempre se les exigiera hacer sus ordenamientos retroactivos, ya que tal requisito incrementaría el costo de los grandes cambios súbitos en la ley elaborada por los jueces.

A veces, la elaboración de leyes por el poder legislativo tenderá a ser más eficiente que la elaboración de leyes por los jueces aun cuando no haya necesidad de cambiar la ley abrupta o drásticamente. A menudo, los hechos de

[12] Véase Louis Kaplow y Steven Shavell, "Why the Legal System Is Less Efficient Than the Income Tax in Redistributing Income", 23 *J. Legal Stud.,* 667 (1994).

[13] Fundamentalmente, cambios en la ley fiscal, que por lo común se hacen levemente retroactivos con el objeto de socavar los esfuerzos que emprenden los contribuyentes para hacer modificaciones en sus ingresos o sus gastos encaminadas a evitar el pago de impuestos. En ocasiones, las leyes que reducen los castigos por delitos se hacen retroactivas, dado que es poco probable que alguien haya depositado su confianza en su severidad.

los casos litigados no ponen de relieve las necesidades que hay para la reforma jurídica. Las personas dedicadas a los negocios pueden (a sus propias expensas) restructurar sus actividades para evadir las normas ineficientes elaboradas por los jueces. Y, en virtud de que a ellos se les exige que den razones para las normas que crean, vacilan en crear normas arbitrarias, aun cuando sucede a menudo que una de éstas es más eficiente que una provista de matices, ya no digamos que un estándar. Compárese una norma que requiere que una apelación se solicite a un plazo de 30 días del veredicto final en el tribunal de instrucción con el estándar que exige que se solicite una apelación dentro de un plazo razonable posteriormente al juicio. Esas normas desempeñan un papel enorme en el sistema jurídico, y la vasta mayoría de ellas son creadas por legislaturas y otros cuerpos hacedores de normas, no a base de opiniones judiciales.

Los "jueces" según este análisis son jueces del derecho común, no del derecho civil. Los jueces del derecho civil, tal como sabemos, tienen menos probabilidades de desempeñar un papel de gran escala en la elaboración de reglas o (un punto relacionado con éste) de ser independientes del gobierno. Los tribunales de derecho civil son, pues, más democráticos que los tribunales de derecho común, en el sentido de que están más adheridos a las legislaturas, las cuales son elegidas, que a los jueces estadunidenses nombrados por designación, quienes ejercen un criterio de amplio alcance, afín al legislativo, no sólo en los casos de derecho común sino también en los constitucionales y en los estatutarios.

XX.3. La teoría económica de la legislación

Una diferencia institucional entre los tribunales y las legislaturas que amerita una consideración separada es la mayor utilización del proceso electoral para seleccionar a los legisladores que para seleccionar a los jueces.[14] Ese proceso crea un mercado de legislación donde los legisladores "venden" protección legislativa a quienes pueden fortalecer sus perspectivas electorales con dinero o votos. Este mercado se caracteriza por agudos problemas de "gorroneo". Alguien que se encuentre dentro del campo de protección de alguna iniciativa de ley se beneficiará de su promulgación independientemente de

[14] Propuesta en George J. Stigler, "The Theory of Economic Regulation", 2 *Bell J. Econ. & Mgmt. Sci.,* 3 (1971). Véanse algunas extensiones en Sam Peltzman, "Toward a More General Theory of Regulation", 19 *J. Law & Econ.,* 211 (1976); Gary S. Becker, "A Theory of Competition among Pressure Groups for Political Influence", 98 *Q. J. Econ.*, 371 (1983); Becker, "Pressure Groups and Political Behavior", en *Capitalism and Democracy: Schumpeter Revisited,* 120 (R. D. Coe y C. K. Wilbur [comps.], 1985). La teoría económica de la legislación, y del comportamiento gubernamental en general, se conoce ahora como de la "elección pública". Véase un excelente resumen en Daniel A. Farber y Philip P. Frickey, *Law and Public Choice: A Critical Introduction* (1991), especialmente la introducción y el capítulo 1; y en cuanto a las críticas sobre la teoría de la elección pública, véase "Symposium on the Theory of Public Choice", 74 *Va. L. Rev.,* 167 (1988).

que haga o no alguna contribución, financiera o de otra clase, para su promulgación. Hay así una analogía estrecha con la cartelización, una analogía reforzada por el hecho de que gran parte de la legislación parece diseñada para facilitar los precios de "cártel" de las empresas reguladas. La analogía nos ayuda a explicar el hecho de que a los consumidores les vaya mal en el proceso legislativo: son demasiado numerosos para organizar un cártel eficaz en apoyo o en oposición a la legislación existente o propuesta.

Los factores que predisponen a una industria hacia la cartelización (discutidos antes, en X.1) son así relevantes para el pronóstico de quiénes triunfarán y quiénes fracasarán en la obtención de la protección legislativa, pero hay diferencias importantes entre el cártel ordinario y la coalición políticamente eficaz. En particular, la escasez de los miembros desempeña un papel menor en la arena legislativa que en la del mercado. Entre menos competidores haya en un mercado, más fácil les será a sus integrantes organizar un cártel privado que tenga pocas probabilidades de ser detectado, de modo que su demanda de protección legislativa podría ser menor que la de un conjunto de competidores similar en otros sentidos pero más numeroso.[15] Segundo, dado que las leyes antimonopólicas no impiden, y no podrían impedir constitucionalmente,[16] que los competidores colaboraran para influir sobre la acción legislativa (por oposición al acuerdo de fijar conjuntamente el precio), los problemas del "gorroneo" se resuelven con mayor facilidad en la arena legislativa que en la del mercado. Y tercero, en la medida en que el número de las empresas que buscan asistencia legislativa se correlaciona positivamente con el número de los individuos empleados por esas empresas o que por lo demás dependen económicamente de ellas, o si los individuos (por ejemplo, los miembros de alguna profesión o de otro grupo ocupacional) están buscando tal asistencia, los grandes números complican el problema del "gorroneo", pero podrían tener un efecto compensatorio al aumentar el poder de votación del grupo.

Estos factores podrían explicar el hecho de que las regulaciones monopólicas sean más comunes en áreas como la agricultura, el trabajo manufacturero y las profesiones —donde la cartelización privada sería generalmente impracticable en vista de los grandes números que tendrían que unirse a cualquier cártel privado para que fuese eficaz— que en las industrias muy concentradas tales como las del acero y el aluminio. Sin embargo, esperaríamos que los grupos de interés exitosos fuesen *relativamente* pequeños y homogéneos, en gran medida como ocurre con las unidades de negociación eficaces en una planta (véase antes XII.2). La escasez de los miembros reduce los cos-

[15] Sin embargo, una calificación importante es que la legislación puede conferir beneficios mayores de los que podría obtener un cártel privado. Por ejemplo, en un mercado de pocos vendedores pero de entrada fácil, la cartelización privada podría ser poco rentable —mientras que la legislación que impidiera nuevas entradas podría ser muy rentable— para tales vendedores.

[16] Eastern Railroad Presidents Conference *vs.* Noerr Motor Freight, Inc., 365 U.S., 127 (1961).

tos de transacción, aumenta el costo del gorroneo (al volver menos probable que haya algo que gorronear; es decir, cada miembro es más importante para el éxito del grupo entre menor sea el número de miembros), aumenta los beneficios de la redistribución y vuelve menos probable la oposición organizada al reducir el costo por oponente. Para entender las dos últimas observaciones, adviértase que si un grupo de 10 toma 20 dólares de otros 100, el costo por transferente es de sólo 20 centavos de dólar y el beneficio por transferido es de 2 dólares; si se invierte el tamaño de los grupos, el costo por transferente será de 2 dólares y el beneficio de sólo 20 centavos de dólar por transferido.

El lector podría preguntarse por qué los grupos políticamente eficaces[17] debieran buscar métodos tan aparentemente ineficientes para transferir riqueza a sí mismos como la limitación de la entrada o la competencia de precios. Una transferencia en efectivo de una suma al contado de una sola vez para la industria, igual al valor presente de las ganancias esperadas de una limitación reguladora de la competencia, daría a la industria los beneficios de la disminución de la competencia sin su costo de peso muerto de la producción reducida (véase antes IX.2). Pero ¿cómo se lograría la recolección de los fondos necesarios para el subsidio en efectivo? La respuesta es la tributación, que como vimos en el capítulo XVIII tiene la misma clase de efectos de sustitución que el monopolio. Más importante aún, el hecho de que la transferencia de riqueza producida por los límites regulatorios a la competencia sea implícita (y, en realidad, con frecuencia sea difícil de estimar) aumenta los costos de información de la oposición a la transferencia, en contraste con un artículo presupuestario explícito: un artículo de grandes proporciones, difícil de pasar desapercibido, puesto que sería una suma al contado equivalente a la totalidad del valor presente que corresponda a la magnanimidad de los legisladores. Y si para reducir su visibilidad la transferencia se paga a lo largo del tiempo, atraerá a nuevos competidores a la industria y se disiparán las ganancias esperadas para las empresas que obtuvieron el generoso regalo.

Todo esto hace que los grupos de interés se vean mal. Pero la verdadera objeción económica no es contra los grupos de interés sino contra el uso del proceso político para la toma de decisiones económicas. Dado que el beneficio de la votación para el individuo es insignificante en cualquier sentido práctico —desvaneciéndose en efecto hasta cerca de cero en cualquier elección fuera de las más locales—, al votante típico no le conviene invertir mucho en aprender acerca de los diversos candidatos o las políticas que defienden.[18]

[17] No siempre tienen que ser grupos de productores. De hecho, gran parte de la regulación es un resultado de las negociaciones entre un grupo de productores y un grupo de consumidores que se ven favorecidos. Véase antes XIII.7.

[18] ¿Resulta sorprendente que sólo cerca de la mitad de la población con derecho de voto se moleste en votar efectivamente? Para un economista, lo sorprendente es que voten tantas personas; pensaríamos que nunca convendría votar mientras el electorado no se redujera a 100 personas o menos. Lo que ha de ocurrir, por consiguiente, es que muchas personas derivan un valor de

Además, los votantes no votan generalmente por las políticas (las iniciativas y los referendos, autorizados en algunos estados, son una excepción) sino por los representantes. Los votantes no pueden obligar a sus representantes a que cumplan sus promesas de campaña, fuera de votar en contra de ellos en la próxima elección (si acaso se presentan para la relección) —no hay ninguna contrapartida del principio fiduciario que reduce los costos de la agencia en otros contextos de principal-agente—. Y, al escoger entre candidatos, están eligiendo los votantes entre paquetes de políticas prometidas. Es posible que ninguno de los paquetes sea particularmente atractivo. En efecto, podrían ser igualmente poco atractivos si, por ejemplo, el votante prefiriera las políticas A y B a las políticas C y D, y el candidato X prometiera A y C, mientras que el candidato Y prometiera B y D.

Los grupos de interés desempeñan un papel esencial en proporcionar información para el electorado y financiar las campañas políticas. Al mismo tiempo, ayudan a superar otro problema del uso de la votación para la asignación de los recursos, a saber: que los votos no reflejan la intensidad de las preferencias como lo hace la "votación" con dólares en un mercado ("cumple con lo que hagas y con tu dinero lo prometido"). Supongamos que un candidato ofrece un paquete de políticas (digamos fuertemente cargadas en favor del antisemitismo, la inflación y los subsidios para las pequeñas empresas) que conferirían un beneficio medio de 1 dólar a cada uno de 200 000 votantes, pero también un costo medio de 1 000 dólares a los restantes 10 000 votantes. Este paquete sería ineficiente, pero, a menos que se le permitiera a la minoría que se le opone formar un grupo de interés para derrotar al candidato, éste bien podría ser elegido.

Asimismo, muchas leyes (la mayoría de las leyes penales, por ejemplo) no son el producto de grupos de interés *estrechos*. Un grupo de interés es menos cohesivo entre más grande sea, pero si los beneficios buscados por el grupo son suficientemente grandes quizá le sea posible todavía superar los problemas de "gorroneo", sobre todo si la oposición es también difusa.[19] Pero adviértase que, una vez que se adopta una ley de interés público, podrían formarse grupos de interés que tengan un interés privado en expandir su alcance, posiblemente más allá del límite eficiente, así como en defenderla si el cambio de las condiciones hace que la ley ya no despierte el interés público. Los funcionarios de los centros de corrección, los policías y los abogados defensores penales tienen un interés pecuniario en el sistema de la justicia penal, aunque ese sistema no sea el producto de las presiones de grupos de interés estrechos.

consumo de la acción de depositar su voto, el cual es análogo al valor que la gente deriva de aplaudir en un concierto, a sabiendas de que ni un solo aplauso de un individuo afectará de modo conspicuo al volumen total de aplausos.

[19] Jürgen Backhaus, "Efficient Statute Law", en *Essays in Law and Economics: Corporations, Accident Prevention and Compensation for Losses,* 23 (Michael Faure y Roger van den Bergh [comps.], 1989).

El papel de los grupos de interés se ve frenado en el sistema judicial, como lo revela claramente la desestimación temprana y firme, por el derecho común, de la proposición de que la competencia es un cuasidelito: una proposición que se encuentra detrás de gran parte de la legislación reguladora. Además de las diferencias procedimentales entre las reglas elaboradas por el sector judicial y el legislativo y señaladas en las secciones precedentes, la elección de jueces, aunque sea ampliamente usada a nivel estatal y local, es generalmente menos partidista, contenciosa y cara que las elecciones para los funcionarios legislativos y ejecutivos (si bien ha habido una modificación al respecto). Por supuesto, al actuar a través de funcionarios elegidos, los grupos de interés pueden influir e influyen sobre las designaciones judiciales. Pero, una vez que el juez toma posesión, está sustancialmente aislado de las presiones de grupos de interés; al nivel federal, donde no hay ningunas elecciones y escasean las oportunidades de promoción, lo está casi por entero.

Un procedimiento para constreñir la influencia de los grupos de interés sobre el proceso legislativo consiste en limitar la cantidad de dinero que se puede gastar en las campañas políticas.[20] Éste es el enfoque utilizado en las Enmiendas de la Ley Federal de Campañas Electorales de 1974.[21] El tope de las contribuciones reduce la ventaja comparativa de que disfruta un grupo que sea capaz de superar los problemas del "gorroneo" al hacer grandes contribuciones a los políticos. Pero hay una gran laguna o resquicio: la Ley permite la clase de contribuciones no pecuniarias que sólo un grupo de interés puede hacer, por ejemplo, los servicios personales de los miembros de un sindicato, organizados y aportados por el sindicato. La propia Ley de 1974 puede ser un ejemplo de la legislación de interés especial, pues el grupo beneficiado son los propios legisladores federales. La ley favorece a quienes ya ocupan cargos en la misma forma en que una limitación a los gastos de la publicidad comercial favorecería a los productores de marcas existentes. Un pro-

[20] Otro consiste en proporcionar subsidios públicos para los gastos de campaña, lo que reduce el costo de realizar una campaña para el político que la encabeza y para sus simpatizantes. Se tiene la esperanza de que, al disminuir la relevancia de los grupos de interés (que se vuelven una fuente de financiamiento de la campaña de segundo orden una vez que están disponibles los fondos públicos), el otorgamiento de subsidios haga más probable que se den soluciones políticas realmente eficientes… pero sólo a condición de que esta práctica vaya aparejada de la imposición de límites más estrictos a las contribuciones hechas por los particulares a las campañas en cuestión, y de que se puedan desarrollar estándares más objetivos para decidir cuál debe ser el monto de los subsidios y cuáles son los criterios de elegibilidad.

[21] Véase Sam Kazman, "The Economics of the 1974 Federal Election Campaign Act Amendments", 25 *Buff. L. Rev.*, 519 (1976). En Buckley *vs.* Valeo, 424 U.S., 1 (1976), por decisión unánime (*per curiam*), la Corte Suprema invalidó, con base en la Primera Enmienda, las limitaciones impuestas por la Ley al gasto total en las campañas, al mismo tiempo que confirmaba las limitaciones impuestas al monto de las contribuciones individuales. Evidentemente, el efecto principal de la Ley modificada por las decisiones judiciales ha sido el de aumentar las contribuciones corporativas. James B. Kau y Paul H. Rubin, *Congressmen, Constituents, and Contributors: An Analysis of Determinants of Roll Call Voting in the House of Representatives,* cap. 8 (1982).

ducto nuevo requiere a menudo gran publicidad a fin de que los clientes potenciales conozcan su existencia y sus ventajas; lo mismo se aplica a una cara nueva en el mercado político. Pero es probable que los efectos de asignación (por oposición a los efectos distributivos) de la limitación de la publicidad política sean menos graves de lo que serían los efectos de la limitación de la publicidad comercial, no sólo debido a las distorsiones creadas por las presiones de grupos de interés, sino también debido al mayor incentivo para la publicidad política fraudulenta que para la publicidad comercial fraudulenta, toda vez que existe una mayor incertidumbre acerca del "producto" ofrecido por el político. La incertidumbre acrecienta la propensión al fraude, como lo hace el hecho de que muchos políticos no planeen permanecer en el mercado por un tiempo indefinido. Por ejemplo, un presidente que busca la relección no afronta ningún castigo en el mercado si deja de cumplir las promesas realizadas en campaña, puesto que si es relegido será, como un presidente saliente a causa de la limitación de dos periodos, alguien con poderes en mengua.[22]

XX.4. La interpretación estatutaria en un mundo de política de grupos de interés

La comprensión convencional que tienen los abogados de la interpretación estatutaria es que el tribunal trata de descubrir y aplicar las intenciones de la legislatura promulgadora. Esto es congruente con ver el proceso legislativo dominado por las negociaciones entre grupos de interés; según esta visión, la promulgación legislativa es una venta negociada y son apropiados los mismos métodos utilizados en la interpretación de los contratos privados ordinarios. Sin embargo, el proceso de descubrimiento de la intención legislativa es más difícil que el del descubrimiento de la intención que se encuentra detrás de un contrato ordinario, debido a la naturaleza plural del órgano promulgador. No puede suponerse automáticamente que las declaraciones de los legisladores individualmente, aun de los comités legislativos, expresan las opiniones de la "mayoría silenciosa" que se necesita para la promulgación. Además, los proponentes de la legislación del grupo de interés podrían ocultar el objetivo verdadero de la legislación a fin de aumentar los costos de información para los oponentes. Pero, por lo menos hasta cierto punto, esta reticencia es contraproducente. Es probable que lo que se oculta al público se oculte a los jueces, lo que los lleva a construir una justificación de interés público que podría moldear la interpretación judicial de tal manera que aminore el impulso redistributivo de la legislación. A veces la interpretación judicial amplifica ese impulso: cuando el juez favorece personalmente la redistribución, o cuando el juez no logra acordar una solución intermedia diseñada para atenuar el em-

[22] Más adelante, en XX.7, discutiremos la reforma del financiamiento de las campañas.

puje redistributivo de la legislación. El segundo factor a menudo se pasa por alto cuando los jueces crean a veces un remedio de daños privado para la violación de un estatuto regulador si creen que el estatuto se aplicaría con mayor eficacia permitiendo que las víctimas de la infracción iniciaran juicios por daños. Puede asistirles la razón en pensar así (véase después XXIV.1), pero no en dar por sentado que se proyectaba hacer cumplir el estatuto con la máxima efectividad. Es posible que los oponentes del estatuto tuviesen fuerza suficiente para imponer una solución negociada o intermedia que limitara la eficacia del estatuto al no autorizar la aplicación rigurosa de éste por la esfera privada. De ser así, la creación judicial de un remedio de daños privado destruirá dicha solución.

Los jueces consultan, a veces, la historia legislativa posterior a la promulgación a fin de aclarar el significado de un estatuto. Ése es un error. Es posible que el equilibrio del poder en la legislatura haya cambiado tras la promulgación del estatuto; la historia legislativa posterior a la promulgación podría ser un esfuerzo para deshacer el trato logrado por los legisladores anteriores.

Ocasionalmente, los jueces utilizarán la política incorporada en un estatuto para precisar el significado de otro, en gran medida como si el primer estatuto fuese un precedente judicial. El peligro es que de esta forma se minará el acuerdo logrado en el segundo, a la vez que se ampliará en efecto el campo del acuerdo logrado en el primer estatuto más allá de su dominio intentado, socavando así igualmente el acuerdo sostenido en ese estatuto previo. Sólo si ambos estatutos fueran de interés público tendría sentido utilizar el primero como una razón para entender el segundo en una forma particular.

La utilidad del análisis económico en la interpretación estatutaria no se limita a recordar a los jueces que vigilen a los grupos de interés. Consciente de que la información es un bien costoso, el economista puede recordarle también al juez que es poco realista suponer, como lo hacen a veces los jueces, que cada palabra de un estatuto tiene una función distintiva al impartir significado al estatuto. Los costos de información podrían conducir a la omisión y la redundancia, a menudo en el mismo estatuto. Además, los términos estatutarios redundantes, justo como cuando levantamos la voz en una conversación telefónica en los casos en que la conexión es mala o repetimos un mensaje para asegurarnos de que ha sido recibido,[23] sirven a la eficiencia superando el ruido existente en el canal de comunicación. De esta manera, la redundancia puede ser deliberada, pues induce a que todo intento por encontrar significados diferentes a los de los términos empleados constituya un error, el cual irá en desdoro del significado que se propone transmitir el estatuto.

En las áreas donde los métodos de interpretación convencionales dejan al juez en duda, es probable que la interpretación de éste refleje sus juicios

[23] Véase Kerr S. S. Co. *vs.* Radio Corp. of America, 157 N.E., 140 (N.Y., 1927), del que se trata en Richard A. Posner, *Cardozo: A Study in Reputation,* 115-116 (1990).

previos. Se ha hecho un llamado a que en casos así el juez se incline a favor de la interpretación que promueva la eficiencia.[24] Esta propuesta subraya el hecho de que en los casos de incertidumbre interpretativa la "interpretación" estatutaria no sea realmente interpretativa, sino antes bien una forma de la elaboración judicial de políticas. Esta posibilidad se explora después, en XX.6.

XX.5. El poder judicial independiente y la política de los grupos de interés

Vimos en el capítulo IV que en el caso de los contratos privados, a menos que el cumplimiento sea simultáneo por ambas partes del contrato o que pueda confiarse en que el deseo de negocios futuros asegure el cumplimiento fiel, las partes contratantes insistirán en que haya un tribunal o un árbitro —un tercero independiente— al que puedan recurrir para el cumplimiento del contrato. Pero no hay ningún mecanismo de cumplimiento legal para la aplicación de sanciones a la legislatura que deje de negociar con un grupo de interés. Por ejemplo, si la industria aeronáutica obtiene del Congreso (como lo hizo en 1938) una legislación destinada a promover los precios monopólicos al mismo tiempo que se impide la entrada de los nuevos competidores a los que atraerían normalmente tales precios, el Congreso promulgador no podrá impedir que un Congreso posterior enmiende la legislación en forma desfavorable para las aerolíneas, o en efecto que la derogue por completo (como finalmente ocurrió… pero sólo después de 45 años). Es cierto que la mala fe del Congreso reduciría el valor presente de la protección legislativa de grupos de interés en el futuro y, por ende, el bienestar futuro de los congresistas. Pero para muchos congresistas, sobre todo aquellos que no esperaban permanecer mucho tiempo en el Congreso, los beneficios de repudiar el arreglo que hizo un Congreso anterior podrían superar a los costos. Cuando es apretada la votación inicial sobre la promulgación de la legislación, el que apenas unos pocos congresistas de la coalición ganadora deserten, a resultas del retiro o la derrota en las urnas, podría conducir a una derogación en la siguiente sesión del Congreso; los congresistas recientemente elegidos no tienen ningún compromiso de respetar los tratos de sus predecesores.

La estabilidad necesaria para los tratos legislativos a largo plazo deriva de: *1)* las reglas de procedimiento de la legislatura y *2)* la independencia del poder judicial. La regla procedimental más importante es el requerimiento de que la legislación (incluida la legislación de enmienda o derogación) sea aprobada por una mayoría de los legisladores votantes. Este requerimiento convierte a la promulgación legislativa en un proceso difícil y tardado, debido a

[24] Como se aconseja en Jonathan R. Macey, "Promoting Public-Regarding Legislation through Statutory Interpretation: An Interest Group Model", 86 *Colum. L. Rev.,* 223 (1986). Véase Einer R. Elhauge, "Does Interest Group Theory Justify More Intrusive Judicial Review?", 101 *Yale L. J.*, 31 (1991).

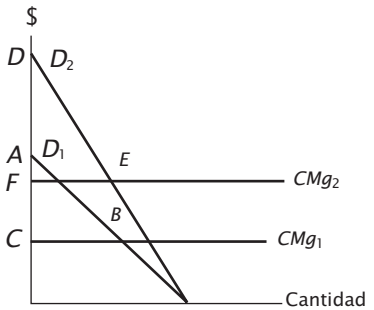
los costos de transacción implicados en la obtención del acuerdo entre gran número de individuos. Una vez aprobado un estatuto, por lo tanto, es improbable que sea sustancialmente alterado o revocado, dada la presión de otros asuntos legislativos.[25] Los impedimentos para legislar dotan a la legislación, una vez promulgada, de cierto grado de durabilidad, lo que aumenta el valor y, por ende, la demanda de legislación.

Demos por supuesto que los impedimentos también reducen la productividad de los gastos dirigidos a obtener la legislación en primer lugar. Sin embargo, bajo supuestos plausibles, el aumento del valor de la legislación superará el aumento de su costo; esto se muestra en la gráfica XX.1. D_1 representa la curva de demanda de legislación de interés especial con el supuesto de que los beneficios de la legislación se limitarán a un solo periodo (es decir, la duración de la legislatura que la promulga). La curva de demanda tiene pendiente negativa porque algunos grupos obtendrán de la legislación protectora mayores beneficios que otros y, por lo tanto, estarán dispuestos a pagar por ella un precio mayor. CMg_1 es el costo en que incurre la legislatura para promulgar tal legislación. Los beneficios netos de la legislación, presumiblemente divididos entre los legisladores y los grupos de interés, son entonces el área *ABC*. Con todo, si los beneficios de la legislación pueden obtenerse durante un periodo mayor que el de la existencia de la legislatura promulgadora, la curva de demanda se convierte en algún múltiplo vertical de D_1 (tal como D_2),[26] y los beneficios netos de la legislación serán mayores (*DEF*) que en el modelo del periodo único, aunque los costos de la promulgación de la legislación sean también mayores (CMg_2).

Alguna legislación es ineficaz sin sustanciales asignaciones anuales de la legislatura para pagar un subsidio periódico o sufragar los gastos de una oficina pública encargada de hacer cumplir el estatuto. La legislación incompleta en este sentido, en el momento de la promulgación, es mucho menos valiosa para sus beneficiarios, quienes podrían tener que "comprar" de nuevo la legislación cada año. Por lo tanto, independientemente del problema de la atracción de la entrada (véase antes XX.3), esperaríamos que la legislación de grupos de interés asumiera por lo general una forma que evite la necesi-

[25] Otras características del proceso legislativo generan también resistencia a la promulgación acelerada de leyes nuevas: el bicameralismo, el sistema de comités, los filibusteros y la tradición de nombrar a los presidentes de comités con base en la antigüedad. El uso de la antigüedad tiende a canalizar las presidencias hacia los tenedores de escaños "seguros": los legisladores cuya permanencia puede esperarse que sea larga y quienes tienen, por lo tanto, mayores probabilidades de respetar el trato legislativo, aunque, dependiendo de la rotación, para el momento en que un legislador asciende por la escalera de la antigüedad hasta la presidencia de un comité podría ser tan anciano que su permanencia esperada en la presidencia fuera breve.

[26] La razón por la que la curva de demanda no es sencillamente D_1 por la cantidad de periodos en los que se espera que la legislación siga teniendo vigencia es que los beneficios futuros deben ser descontados al valor presente, a fin de determinar hasta qué grado posee valor la legislación para quienes la apoyan.



GRÁFICA XX.1

dad de aportaciones anuales de carácter sustancial. Un ejemplo importante es el de la legislación que crea agencias reguladoras que utilizan el poder sobre las tasas y la entrada para redistribuir la riqueza: los presupuestos anuales de estas agencias son triviales en relación con las redistribuciones que efectúan. Y, cuando se utilizan subsidios directos, su financiamiento se hace a menudo independientemente de la acción legislativa posterior mediante el instrumento del impuesto destinado a fines particulares, como ocurre en los programas de carreteras interestatales y de seguridad social.

La experiencia de la Prohibición (de la producción y venta de bebidas alcohólicas) ilustra el problema de una legislación que requiere sustanciales aportaciones anuales para conservar su eficacia. Los partidarios de la Prohibición pudieron obtener una enmienda constitucional, lo que es normalmente una forma de legislación particularmente durable. Pero la prohibición de la venta de bebidas alcohólicas requería de un esfuerzo masivo de cumplimiento de la ley que los congresos subsecuentes no estuvieron dispuestos a sostener con suficiente dinero, de modo que la enmienda constitucional fue efectivamente nulificada, y se derogó en 1933 tras haber estado vigente durante sólo 13 años.

Ninguna legislación es de cumplimiento completamente automático. Si las personas sujetas a una ley se niegan a obedecerla, se requerirá el recurso a los tribunales. Un poder judicial que fuese dócil a los miembros de la legislatura actual podría anular efectivamente, mediante una interpretación, la legislación promulgada en una sesión anterior de la legislatura. Es menos probable que los jueces hagan esto si los periodos de la actuación judicial los hacen independientes de los deseos de los legisladores actuales.[27]

Sin embargo, dado que un poder judicial independiente es una fuente de costos y de beneficios para la legislatura, esperaríamos que se diera al poder judicial una independencia menor entre más breve sea la duración esperada de la legislación de interés especial. Ésta podría ser una razón del hecho de

[27] Véanse algunas pruebas empíricas de que las legislaturas aprecian realmente la independencia judicial en Gary M. Anderson, William F. Shughart II y Robert D. Tollison, "On the Incentives of Judges to Enforce Legislative Wealth Transfers", 32 *J. Law & Econ.*, 215 (1989).

que, a medida que bajamos del nivel gubernamental federal al estatal y el local, encontremos periodos más breves para los jueces y una mayor dependencia de la elección antes que de la designación como método para seleccionarlos. Entre más confinada o local sea la jurisdicción de una legislatura, menor será su campo para la redistribución de la riqueza. Por lo tanto, si de todos modos no es probable que los grupos de interés busquen acuerdos duraderos de las legislaturas estatales y locales, las ramas políticas estarán menos dispuestas a pagar el precio de un poder judicial independiente.

XX.6. ¿QUÉ MAXIMIZAN LOS JUECES?[28]

Hemos visto que: *1)* cuando los jueces son los elaboradores del derecho sustantivo, las normas del derecho tenderán a ser congruentes con los dictados de la eficiencia, pero: *2)* cuando los jueces están aplicando los estatutos lo harán de acuerdo con los términos del "trato" original entre la legislatura que los promulgó y los beneficiarios de la legislación. ¿Pueden conciliarse estas proposiciones entre sí y con el supuesto del economista de que las personas maximizan racionalmente su interés propio? Puede darse por sentado que los jueces, siendo personas, tratan de maximizar su utilidad, pero las normas que rigen sobre el proceso judicial han sido diseñadas para impedir que el juez reciba un pago monetario por decidir un caso en una forma particular y asimismo para minimizar la influencia de las lealtades personales y de clase, la amistad, la ideología, el temperamento y los grupos de interés políticamente eficaces sobre sus decisiones. La eficacia de estas normas de aislamiento ha sido cuestionada. Por ejemplo, en la literatura añeja se sostiene a veces que el juez que es propietario de tierras decidirá en favor de los terratenientes, el juez que camina al trabajo decidirá en favor de los peatones, el juez que era antes un abogado corporativo decidirá en favor de las corporaciones. Pero cuando un fallo particular promovería los intereses de un grupo al que ya no pertenece el juez (nuestro último ejemplo), el interés propio del juez no se beneficia seleccionando ese fallo, aunque su experiencia anterior, su ideología o sus señas de identidad en materia étnica o en otros renglones socioculturales (en suma, sus juicios previos) podrían llevarlo a evaluar los méritos del caso de un modo diferente al de los jueces de antecedentes distintos.

Es menos probable que pese significativamente sobre las decisiones del juez un aumento de su ingreso por fallar en favor de un grupo amplio, como el de los peatones o el de los propietarios de casas, al que todavía pertenece; el beneficio financiero personal será de ordinario tan trivial que se verá casi siempre superado por los castigos, por moderados que sean (la crítica profe-

[28] Véase Lee Epstein, William M. Landes y Richard A. Posner, *The Behavior of Federal Judges: A Theoretical and Empirical Study* (2013).

sional, la revocación por un tribunal superior, etc.), por fallar un caso en una forma percibida como poco razonable o sesgada. No es sorprendente, así, que los intentos por conectar las políticas y los fallos judiciales con los intereses económicos personales de los jueces hayan fracasado[29] y hayan desaparecido en gran medida.

Con esto no se niega la importancia del interés económico propio de los jueces en la explicación del comportamiento judicial o el valor de un modelo de maximización de la utilidad para el comportamiento judicial,[30] sino sólo el hecho de que el dinero sea el elemento de mayor envergadura en las funciones de utilidad de la mayoría de los jueces,[31] en especial al referirse a las funciones de utilidad de los jueces federales, que tienen una titularidad de por vida y rara vez renuncian para optar por la práctica profesional, donde podrían hacer más, a menudo mucho más dinero. Pero todo esto significa que al analizar el comportamiento judicial en términos de interés propio debemos enfocarnos en los elementos no monetarios de una función de utilidad.

Para muchos jueces (uno esperaría que para todos), un elemento importante es el de acatar las normas internalizadas de imparcialidad, neutralidad y, en síntesis, el comportamiento judicial respetado, esperado. No obstante, otros elementos que se encuentran entre las funciones de utilidad de muchos jueces acusan el interés propio en el sentido normal de egoísmo. Uno es la preferencia por el ocio, o la aversión al esfuerzo (que no son la misma cosa, ya que el esfuerzo es una función de la intensidad así como del tiempo). En virtud de que los jueces no reciben aumentos como remuneración por su buen trabajo, no sería de esperarse que un aumento de los sueldos judiciales provocara un trabajo más arduo de los jueces. En efecto, si se supone que el ocio es un bien superior (un bien cuya demanda sube con el ingreso), un aumento de los sueldos judiciales podría hacer que los jueces trabajaran *menos* arduamente. Pero ése no sería el único efecto. El aumento salarial reduciría la rotación judicial y aumentaría el número de los candidatos a jueces. También podría cambiar la mezcla de candidatos en favor de quienes asignan mayor

[29] Roscoe Pound, "The Economic Interpretation and the Law of Torts", 53 *Harv. L. Rev.,* 365 (1940); Orley Ashenfelter, Theodore Eisenberg y Stewart J. Schwab, "Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes", 24 *J. Legal Stud.,* 257 (1995). En el capítulo VIII examinamos algunos esfuerzos para explicar los fallos judiciales sobre la base del sesgo de clase.

[30] Véase por ejemplo, antes Epstein *et al.*, nota 28; Richard A. Posner, *Overcoming Law,* cap. 3 (1995); Ahmed E. Taha, "Publish or Paris? Evidence of How Judges Allocate Their Time", 6 *Am. Law & Econ. Rev.,* 1 (2004).

[31] No obstante, existen evidencias de que, entre más generosa sea la pensión judicial, más temprana será la edad en la que se jubilan los jueces (o bien optan por la semijubilación, según se permite en el sistema judicial federal). Albert Yoon, "Pensions, Politics, and Judicial Tenure: An Empirical Study of Federal Judges, 1869-2002", 8 *Am. Law & Econ. Rev.,* 143 (2006). El efecto se intensifica conforme uno se desplaza a través de la jerarquía judicial, ya que a medida que se atenúan las perspectivas de obtener más adelante una promoción disminuyen los incentivos para quedarse en el puesto.

importancia al ingreso que el candidato típico a un puesto de juez cuando los sueldos judiciales son menores. Los salarios bajos tienen un efecto de separar el grano de la paja, y a veces dejan en fila aparte a los aspirantes cuyo motivo es exclusivamente financiero… no la clase de personas que querríamos tener como jueces.

Si es probable que el aumento de los sueldos de los trabajadores judiciales haga que los jueces trabajen menos arduamente, esto podría explicar el hecho de que el Congreso combinara en 1991 un aumento salarial sustancial para los jueces federales con una intensificación sustancial de las restricciones impuestas al ingreso obtenido por los jueces por tener otros trabajos al mismo tiempo (principalmente en la enseñanza). Tales restricciones disminuyen el costo de oportunidad del tiempo dedicado a la actividad judicial. Sin embargo, al mismo tiempo disminuyen el costo de oportunidad del ocio, de modo que no hay ninguna seguridad de que tales restricciones conducirán a una producción judicial mayor.

No podemos tener la certeza de que un aumento en el número de casos bajo su jurisdicción produzca un trabajo más intenso por parte de los jueces. Ello no se debe a que los jueces estén trabajando a su máxima capacidad, sino a que pueden responder a dicho aumento, sin sacrificar su ocio, invirtiendo menos horas en cada caso o simplemente tomándose más tiempo para decidir sobre cada uno, dejando que se acumulen las pilas de expedientes. De la misma forma, es poco probable que al agregar jueces a un tribunal se genere un notable incremento en la productividad, puesto que algunos de ellos, tras percibir que hay menos presión para que por su cuenta den término a equis casos, no trabajarán con el mismo ahínco.[32]

Una manera en que la aversión al esfuerzo se manifiesta a menudo es la de la renuencia a disentir, ya que poner por escrito un voto particular de discrepancia (o incluso tan sólo editar el borrador de un disenso redactado por un asistente jurídico) toma tiempo, y un disenso no reduce el número de opiniones de mayoría que el juez tiene el deber de plasmar en el papel y también causa irritación entre los jueces que integran la mayoría. Esa irritación, incluso si no conduce a desquitarse del disidente, socava el espíritu de camaradería, generando un costo no pecuniario (condiciones laborales menos atractivas).

La aversión al disenso podría explicar por qué un panel de apelación federal en el que los jueces no fueron nombrados por un presidente del mismo partido probablemente decida acerca de un caso políticamente controvertido, como el de la discriminación de género o el aborto, de forma distinta a como lo hará un panel cuyos jueces han sido designados por un presidente del mismo partido.[33] Específicamente, los estudios encuentran evidencias de que existe una "moderación [ideológica] por parte de la posición mayoritaria cuando

[32] Véase Michael Beenstock y Yoel Haitovsky, "Does the Appointment of Judges Increase the Output of the Judiciary?", 24 *Int'l Rev. Law & Econ.*, 351 (2004).

[33] Véase por ejemplo Cass R. Sunstein *et al.*, *Are Judges Political? An Empirical Analysis of the*

hay un panel de dos en vez de tres jueces designados por un presidente del mismo partido".[34] Una posible explicación es que el hombre que se sale del molde actúa como un soplón. Otra es que este sujeto puede aportar a las deliberaciones del panel ideas penetrantes que, a los otros jueces, fincados en supuestos elaborados de antemano y presumiblemente diferentes, pueden haberles pasado inadvertidas. Con todo, un factor mayor pueden ser las diferencias entre los miembros del panel respecto a la intensidad con la que manifiestan su preferencia por un fallo en particular, como se puede inferir por la distancia ideológica entre el juez disidente y la mayoría de sus colegas respecto al caso en particular —entre mayor sea la distancia, más probable será la disensión… junto con la aversión a ésta—. Si un juez expresa con vehemencia su sentir de que el caso debiera decidirse de un modo y no de otro, mientras que los otros dos jueces, aunque se inclinen a votar de ese otro modo, no se expresan con la misma vehemencia, uno de ellos puede decidir alinearse con el tercero para evitar que cunda la mala voluntad, y quizás esperando que haya consideraciones recíprocas para algún caso futuro en el que se incline fuertemente hacia un lado y los otros dos jueces no. Una vez que un juez vira hacia el ángulo del juez disidente, el juez restante probablemente hará lo mismo por razones similares o a causa de su aversión a expresar su disenso.

El desacuerdo ideológico no se parece al desacuerdo sobre el mejor medio para conseguir un fin acordado por las partes porque los contendientes ideológicos rara vez discuten a partir de premisas compartidas. Un liberal incluido en un panel donde hay dos conservadores difícilmente sacará a flote hechos o argumentos a fin de cambiar la ideología de sus colegas, y viceversa. No obstante, si su sentir acerca de cómo ha de decidirse el caso es más intenso que el de los otros jueces, esto implica que él derivaría mayores beneficios de una decisión del caso ajustada a su parecer que los que ellos obtendrían si se decidiera según sus preferencias, y, por lo tanto, implica también que él estaría dispuesto a incurrir en costos más grandes para salirse con la suya, por ejemplo, al poner por escrito un disenso. Así pues, su amenaza de disentir es una amenaza creíble de imponer costos a sus colegas (los costos derivados de la aversión que éstos sienten por el disenso) en el caso de que ellos se rehúsen a seguir sus preferencias. Si tales costos exceden a los beneficios de un fallo favorable para al menos uno de sus colegas, toda vez que a ese colega no le incumbe tanto el fallo, éste bien puede ceder.

Otro elemento en la función de utilidad de algunos jueces es la ambición; es bien sabido que los jueces se afanan en pescar una promoción, por ejemplo, imponiendo largas sentencias a los delincuentes,[35] que son claramente

Federal Judiciary (2006); Richard L. Revesz, "Environmental Regulation, Ideology, and the D.C. Circuit", 83 *Va. L. Rev.,* 1717 (1997).

[34] Sunstein *et al.*, nota 33 antes, en 20-21.

[35] Véase Epstein *et al.*, nota 28 antes, cap. 8.

un segmento impopular de la sociedad. Otro componente plausible de la función de utilidad judicial es el prestigio. Consideremos ahora el efecto, sobre la calidad del poder judicial federal, de un aumento sustancial del número de jueces federales, como se pide de tiempo en tiempo. Al aumentar la razón de los jueces a los juicios, tal aumento reduciría las cargas de trabajo judiciales, volviendo un puesto de juez más atractivo para una persona que valuara el ocio grandemente. Al mismo tiempo, al aumentar la razón de los jueces a los abogados o a los ciudadanos, un aumento del número de jueces reduciría el prestigio de ser un juez. Por lo tanto, podría esperarse que tal aumento alterara el conjunto de candidatos a jueces en contra de los maximizadores del prestigio y en favor de los maximizadores del ocio.

Una parte de la función de utilidad de muchos jueces es un anhelo, natural en un sistema en el que los jueces tienen poder por incomparecencia para elaborar reglas —debido a las muchas lagunas y ambigüedades habidas en los estatutos producidos por legislaturas poco afectas a la disciplina—, de influir en las políticas públicas. Cuando ejercen esa influencia bajo la apariencia de realizar una interpretación estatutaria, están jugando un juego, en el sentido de la teoría de juegos, con la legislatura. Si los jueces de este carácter dominan en la Corte Suprema, podremos pronosticar que es improbable que una decisión de la Corte que interprete un estatuto federal sea derogada por el Congreso durante el mismo periodo legislativo en el que se tome la decisión. La Corte no ganaría nada con un acto inútil y, por lo tanto, no adoptaría una interpretación que sería anulada pronto. Pero la Corte no puede prever perfectamente el futuro, de modo que no es sorprendente que sus decisiones interpretativas sean ocasionalmente revocadas (para el futuro: el Congreso no puede anular las decisiones judiciales, pero puede anular para el futuro la norma formulada en una decisión) en un periodo posterior del Congreso. Es posible que un cambio imprevisto en la composición de sus miembros haya alterado la línea política percibida por la Corte cuando tomó la decisión.[36] El mismo análisis nos ayuda a explicar —algo que es analíticamente lo mismo— la extrema sensibilidad de los jueces al hecho de ser revocados por un tribunal superior; la revocación elimina el efecto de la decisión del juez sobre las partes del caso inmediato y sobre otras, similarmente situadas, cuyo comportamiento podría verse influido por la regla declarada por el juez.

La decisión de acuerdo con el tenor original de la legislación puede explicarse en términos similares. A menos que los jueces decidieran cuestiones de la interpretación estatutaria en ese sentido, la independencia del poder judicial dejaría de desempeñar una función esencial en un sistema político creado en un grado considerable con base en la interacción de grupos de interés y podría verse reducido en consecuencia por la legislatura, con una pérdida

[36] Rafael Gely y Pablo T. Spiller, "A Rational Choice Theory of Supreme Court Statutory Decisions with Applications to the 'State Farms' and 'Grove City' Cases", 6 *J. Law, Econ. & Organization,* 263 (1990).

concomitante de poder judicial.[37] Aun así, los jueces "originalistas", que están políticamente enterados, afrontan un conflicto.[38] Si interpretan un estatuto estrictamente de acuerdo con el entendimiento de los legisladores, y el equilibrio político ha cambiado en la legislatura, su interpretación podría conducir a una revocación rápida del estatuto por la legislatura actual. Si en cambio los jueces dan al estatuto una interpretación que preserve su sustancia, pero la inclinan un poco hacia la política preferida de los legisladores actuales, la legislatura, con su agenda recargada, podría decidir que el estatuto se quede como está. Un enfoque orientado hacia las políticas, políticamente consciente, interpretativamente flexible, podría producir a final de cuentas un resultado más cercano a la intención legislativa original que un enfoque originalista.

La explicación del otro dato con el que empezamos, el contenido económico implícito del derecho común, podría parecer clara en lo tocante a las áreas —la de los contratos principalmente, pero también grandes tramos del derecho de propiedad y el de los cuasidelitos— donde los costos de transacción son bajos. En tales áreas, las reglas de derecho ineficientes serán nulificadas por el acuerdo expreso de las partes, mientras que el persistente desafío judicial contra la lógica económica inducirá a las partes contratantes a utilizar en cambio el arbitraje privado para la resolución de las disputas contractuales. En las áreas del derecho común donde no existe ninguna relación antecedente entre los litigantes, como el derecho penal y la mayoría de los cuasidelitos, los tribunales no están sujetos a las mismas restricciones competitivas. En realidad, empero, puede sostenerse que hay *más* presión en favor de las reglas eficientes en las áreas de altos costos de transacción, ya que en las áreas de bajos costos de transacción las partes podrían contratar para eludir una regla ineficiente con tanta facilidad que no se plantearía ningún juicio para cuestionar dicha regla, de modo que no habría ocasión para someterla a revisión.

No debemos suponer, y esto es importante, que todos los jueces, incluso los del mismo tribunal, tienen funciones de utilidad idénticas. Algunos pueden ser activos elaboradores de políticas; otros pueden preferir una combinación de ocio con una reputación profesional sólida y mostrar un afán de eludir la controversia. El elaborador de políticas tendrá un impacto mayor sobre la ley, mientras que la adhesión al precedente sirve a los fines del ocio y la reputación y será la política de los jueces con esas funciones de utilidad. Esto ayuda a explicar que el derecho común pueda ser eficiente, aunque la mayoría de los jueces sean completamente indiferentes a la eficiencia. Si los

[37] En el siguiente capítulo se discute hasta qué punto es factible que esa cuestión les concierna a los jueces, dados los problemas del "gorroneo" (resulta improbable que el comportamiento de un solo juez repercuta sustancialmente en la práctica de la decisión de acuerdo al precedente). Véase XXI.2 después.

[38] Véase John A. Ferejohn y Barry R. Weingast, "A Positive Theory of Statutory Interpretation", 12 *Int'l Rev. Law & Econ.,* 263 (1992).

casos no controlados por el precedente requieren juicios en materia de política y de tipo legislativo que el juez elaborador de políticas tenderá a basar en intuiciones económicas, las decisiones de tales juicios llevarán la huella de la lógica económica. Las decisiones que las sigan basadas en el precedente llevarán también esa huella, aunque los jueces que tomen estas decisiones no conciban su papel de elaboradores de políticas.

La referencia al juicio político u orientado hacia políticas nos remonta al espectro de los costos judiciales de agencia. Resulta factible que éstos sean altos en un sistema donde los jueces están en gran medida bien resguardados de la práctica de ejercer control ya sea a través de la zanahoria o del palo. La amenaza de la intervención legislativa si los jueces van "demasiado lejos" establece algunos límites externos a la discrecionalidad judicial. Por añadidura, los candidatos judiciales federales son auscultados cuidadosamente, lo que se halla en consonancia con el principio general del empleo de que un empleador que no puede evaluar directamente la producción o productividad de un empleado (o actuar con base en su evaluación) tenderá en cambio a fundar la contratación, el salario, la promoción y las decisiones de despido en las aportaciones observables, es decir, en las señales *ex ante* de calidad. De manera que uno espera —y uno encuentra— que, entre más segura sea la titularidad de los jueces en un régimen jurídico en particular, y por lo tanto resulte más difícil controlar su conducta, más cuidadosa será la auscultación de su perfil para el puesto por parte de las autoridades designadas. Asimismo, dando por sentado que por lo general las personas manifiestan preferencias estables y que el comportamiento connota un fuerte elemento habitual, entre más edad tenga un individuo cuando es nombrado para un trabajo, más predecible será su desempeño en él; ésta es una razón de que, entre más elevada sea la posición del tribunal, más avanzada será la edad promedio de los jueces que trabajan en él.

Los costos judiciales de agencia deben distinguirse de la actividad de emitir juicios políticos. Podrá haber una gran cantidad de estudios que confirman que el partido político del presidente recién nombrado es un buen predictivo de la naturaleza de los votos que formule el juez en una amplia variedad de casos, pero eso no prueba sino que en el derecho estadunidense existe un área dilatada, y aún no demarcada por cotos, en la que los materiales legales convencionales no dictan la obtención del fallo y el juez se ve impelido a hacer un juicio sobre políticas que, en su calidad de tal, sufrirá inevitablemente la influencia de preferencias políticas o ideológicas. La metodología formalista que, según se supone, permite a los jueces decidir sobre los casos, por indeterminada que parezca al estar desprovista de referencias a las ideas que tenga el juez sobre las políticas públicas de carácter sólido, "no es por mucho tan científica como lo pretende ser […] Los jueces debieran rechazar las conceptualizaciones mojigatas […] a la hora de contemplar los papeles que ellos cumplen, y asumir en cambio los inevitables roles que cumplen como

hacedores de políticas *de facto*"[39] en muchos casos, siendo éstos generalmente los más importantes. Surge un problema de costos judiciales de agencia sólo cuando la laxitud del control ejercido por el principal sobre el juez-agente le da a éste la oportunidad de tomar decisiones inducidas por una preferencia que es demasiado personal, demasiado peculiar para ser considerada legítima.

El concepto de función de utilidad judicial puede explicar algunos rasgos enigmáticos de nuestro sistema judicial, como la enorme disparidad entre las tasas de disenso de los tribunales de apelaciones a escala federal y la Corte Suprema: aproximadamente 8% frente a 60%. En parte, esto refleja el hecho de que los jueces del tribunal de apelaciones se sientan por lo común en un panel integrado por tres jueces, mientras que en la Corte Suprema (excepto en casos de recusación) los miembros del panel que atienden casos suman nueve. Aun si todos los jueces muestran la misma propensión a disentir, habrá más disensos entre más jueces escuchen un caso. Pero no siete y media veces más. La aversión al esfuerzo, la aversión al conflicto y la preferencia en materia de elaboración de políticas pueden explicar el resto de la diferencia. Los tribunales de apelaciones a nivel federal tienen cargas de trabajo más pesadas que la Corte Suprema, de modo que el costo por el esfuerzo de disentir es más grande. Al mismo tiempo, el beneficio en cuanto a elaboración de políticas que acarrea el disenso es menor porque los tribunales de apelaciones rara vez anulan a sus propios precedentes. Esto se relaciona con la carga de trabajo: la decisión tomada de acuerdo con el precedente economiza el tiempo de los jueces, en virtud de que significa seguir una decisión previa en vez de sopesar los méritos de la apelación partiendo de su base hasta llegar a la cúspide. La Corte Suprema se halla menos limitada por sus precedentes no sólo porque tiene más tiempo en sus manos sino también porque el impacto que tiene sobre las políticas es mucho mayor que en el caso de los tribunales de apelaciones, de manera que los jueces son menos renuentes a arrumbar en un rincón un precedente que está bloqueando su fallo a favor de su política preferida. Así pues, un juez disidente puede tener una esperanza fundada en la realidad de que la decisión de la que él disiente será algún día anulada, y entonces su disensión devendrá en la ley.

XX.7. La teoría económica de la democracia

Hasta ahora, el proceso político democrático se ha descrito en este capítulo casi enteramente en términos de los legisladores, considerados como agentes de grupos de interés, y de los jueces. Una teoría del proceso político basada en los grupos de interés está incompleta. Debe insertarse en una teoría económica más amplia del proceso democrático.

[39] Edward Cantu, "Posner's Pragmatism and the Turn toward Fidelity", 16 *Lewis & Clark L. Rev.,* 69, 76 (2012).

Tal teoría, de ordinario llamada democracia de élite o competitiva, fue propuesta por Joseph Schumpeter.[40] En ella, los miembros de una élite (la clase política o gobernante) compiten por los cargos y el poder, mientras que el público votante funciona, por lo menos durante la mayor parte del tiempo, como poco más que una audiencia cuyo aplauso (los votos) determina cuáles competidores pertenecientes a la élite prevalecerán. "El método democrático es el arreglo institucional para la toma de decisiones políticas donde los individuos adquieren el poder de decidir por medio de una lucha competitiva en pos del voto de la gente."[41] Así la competencia política a la competencia económica, tomando los votos el lugar de las ventas y tomando el poder de las ganancias, y, con los dos lados del mercado claramente diferenciados: los vendedores representan el lado activo y los compradores el lado pasivo. En los mercados económicos, como en los mercados políticos, el comprador no diseña el producto, sino que escoge de un menú que le presentan los vendedores. Los beneficios de la división del trabajo se cosechan al constituir un cuerpo de especialistas en gobernar, a saber, los diputados elegidos y otros funcionarios, liberando así el tiempo del resto de la ciudadanía para que se especialice en otras actividades.

La separación resultante que se da en materia de perspectivas y conocimientos entre gobernantes y gobernados es la contrapartida de la separación existente en los mercados económicos, en materia de perspectivas y conocimientos, entre vendedores y consumidores. Pero mientras que los compradores de los mercados económicos tienen fuertes incentivos financieros o de otro modo fincados en el interés propio para escoger con cuidado entre vendedores rivales, y de ordinario tienen un conocimiento suficiente para determinar cuál vendedor está haciendo la mejor oferta, el ciudadano —el comprador en el mercado político— parece no tener ningún incentivo para votar en absoluto, ya no digamos para invertir esfuerzos en saber cuál candidato hace la mejor oferta, ya que un solo voto no determinará la elección. Y se le pide que compre un "producto" —el candidato y las políticas probables del candidato— cuyo valor es casi imposible de determinar, aunque el votante invierta irracionalmente gran cantidad de tiempo y esfuerzo en el estudio de los candidatos y las cuestiones en juego.

También se puede expresar la diferencia existente entre el mercado económico y el mercado político diciendo que los "bienes" de este último no tienen un precio. El precio es una señal barata y correcta; compacta una información enorme. Los consumidores de altos ingresos pueden usar el precio como un índice de la calidad; los consumidores de bajos ingresos utilizan el precio como una guía hacia los bienes más baratos que se ajusten a una calidad mínima. No hay ninguna señal correspondiente que guíe al votante.

[40] Véase Joseph A. Schumpeter, *Capitalism, Socialism, and Democracy*, caps. 22-23 (1942). La teoría esencial se presenta en menos de cinco páginas. Véase *ibid.,* 269-273.

[41] *Ibid*., 269.

Además, hay en los mercados políticos mayores economías de escala que en los mercados económicos y, por lo tanto, tendencias monopólicas más fuertes. La razón es la existencia de graves problemas de información que afectan a los mercados políticos. "La escala de la actividad política es grande […] porque muchos cargos y dependencias [políticos] ligan a numerosas actividades […] A un electorado con una cantidad limitada de información política le resulta más fácil colocar a una persona a cargo de muchas actividades que escoger a una persona para cada actividad."[42] En consecuencia, los únicos partidos políticos importantes en los Estados Unidos son nacionales, y sólo hay dos, lo que vuelve duopólica la competencia política. Como sabemos, los duopolistas se coluden a menudo en lugar de competir vigorosamente entre sí.

Pero la perspectiva económica ofrece también cierta tranquilidad en lo tocante al proceso democrático. Consideremos la apatía del votante. Si observamos el lado del comprador de un mercado de consumo de buen funcionamiento, vemos lo que podría llamarse una apatía, pero también podría llamarse una aceptación gustosa. Los compradores no necesitan ser compradores alertas, asiduos, cuando creen correctamente que el mercado en el que están comprando es competitivo; tienen una razonable seguridad de que los productos que les ofrece el mercado serán de precio y calidad comparables y satisfactorios. Los vendedores de un mercado competitivo, aunque compiten entre sí vigorosamente, ofrecen productos comparables en precio y calidad, porque de otro modo no sería competitivo el mercado, sino monopolizado. Lo mismo ocurre en el mercado político, y, de este modo, el carácter a menudo similar a Tararí y Tarará (los personajes de la novela *Alicia a través del espejo*, de Lewis Carroll) que presenta la competencia de los grandes partidos de los Estados Unidos no indica que la competencia no esté funcionando. (Por el contrario, el carácter altamente polarizado de la competencia política en los Estados Unidos es señal de que el mercado está fallando.) Asimismo, en los mercados políticos, como en los económicos, los "consumidores" relativamente desinformados, esto es, los votantes, pueden utilizar y en verdad utilizan atajos para cubrir sus déficits de información, como el de inferir la conveniencia de un candidato por la identidad de sus seguidores y sus oponentes.[43]

El hecho de que la mayoría de los consumidores no busquen activamente la mayoría de los productos en cualquier momento dado sirve para aminorar los cambios de la demanda y la oferta, evitando las saturaciones y escaseces que se producirían si todo el público consumidor se precipitara al unísono

[42] Gary S. Becker, "Competition and Democracy", 1 *J. Law & Econ*., 105, 108 (1958). El artículo de Becker, junto con Anthony Downs, *An Economic Theory of Democracy* (1957), refina la teoría de la democracia de Schumpeter.

[43] Véase, por ejemplo, Donald A. Wittman, *The Myth of Democratic Failure: Why Political Institutions Are Efficient*, caps. 2-3, 5 (1995); Samuel L. Popkin, *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns* (1991). *Cf.* Paul Milgrom y John Roberts, "Relying on the Information of Interested Parties", 17 *Rand J. Econ.*, 18 (1986).

sobre un producto nuevo. Los precios cambiarían hasta que se restableciera el equilibrio, pero el cambio de los precios no deja de tener un costo y mientras tanto habría incertidumbre, filas, desempleo y otros problemas económicos. El interés por la estabilidad es mayor aún en la esfera política debido a las consecuencias potencialmente desastrosas de los cambios repentinos y profundos en la gobernanza política. Imaginemos lo que ocurriría si todos los ciudadanos fuesen estudiosos ávidos de la teoría política y se quedaran deslumbrados ante la teoría política radical, generosamente desinteresada de un empresario político carismático y, en consecuencia, eligieran funcionarios que quisieran alterar el curso de la nación en 180 grados. Un elemento que mantiene ese proyecto lejos de nuestro alcance es que la mayoría de los ciudadanos no se interesan por lo que sea mejor para la nación o para el mundo en algún sentido, sino por lo que es mejor desde el punto de vista de su propio interés. Una preocupación por sí mismo genera resistencia al cambio social radical.

La teoría económica de la democracia actúa como un correctivo para la teoría de la elección pública, la que destaca la manera como el gobierno no acierta a gobernar en aras del interés público. La teoría democrática reconoce este punto, pero señala que la falla es el precio de la democracia, o quizá simplemente el precio del gobierno, ya que los regímenes no democráticos casi siempre son menos eficientes que los democráticos. Las fricciones del gobierno son costosas, pero lo mismo ocurre con la transportación. Aceptamos la necesidad de la transportación y, por ende, los costos correspondientes, y debiéramos hacer lo mismo en lo tocante al gobierno.

La controversia sobre la "democracia corporativa" provee un paralelo sugerente con la teoría económica de la democracia. Vimos antes, en XV.8, la preocupación por separar la propiedad y el control en la corporación moderna. Los accionistas de una corporación pública grande son sus propietarios nominales, correspondientes al electorado en el sistema político, y eligen a los directores, quienes a su vez designan a los administradores. Sin embargo, como los políticos de la teoría de la democracia política de Schumpeter, los directores y los administradores de la corporación son los verdaderos "gobernantes" y tienen sus propios intereses, los que pueden divergir de los intereses de los accionistas. La Comisión de Bolsa y Valores ha tratado de estimular una participación mayor de los accionistas en la administración corporativa exigiendo que ésta incluya las propuestas de los accionistas en los resúmenes que se distribuyen a los accionistas de una corporación antes de su asamblea anual.[44] Este intento ha fracasado.[45] Como los votantes, la mayoría de los accionistas de corporaciones públicas tienen sólo un pequeño interés en las corporaciones cuyas acciones poseen: demasiado pequeño para otorgarles un

[44] Véase la Regla SEC 14a-8.

[45] Frank H. Easterbrook y Daniel R. Fischel, *The Economic Structure of Corporate Law*, 85 (1991).

incentivo para invertir tiempo y esfuerzo considerables en el monitoreo del desempeño de los administradores de la corporación. Aun así, la junta directiva, elegida por los accionistas, tiene el poder de despedir a los administradores, y la existencia de este poder provee incentivos para que los administradores se desempeñen bien (o por lo menos no demasiado mal) y permite que los administradores sean sustituidos cuando su área fracasa desastrosamente. El papel del electorado en el mercado político es similar.

Asimismo, el sufragio universal posee valor como un método para reducir la búsqueda de rentas políticas, porque los individuos que carecen de representación son una presa natural para los buscadores de rentas que tienen sus manos en las palancas del poder gubernamental. Además, entre mayor sea el electorado, más difícil resulta que los buscadores de rentas potenciales forjen coaliciones electorales para la explotación de las minorías electorales, porque los costos de organización aumentan con el número de las personas que deben ser atraídas a la organización para que ésta resulte eficaz. Antes se solía temer que la democracia alentara el pillaje a costa de los ricos, simplemente porque son una minoría; pero los costos que representa para la mayoría matar a la gallina de los huevos de oro desalientan esta forma de la explotación. Entre más se grave a los ricos, menos ingreso gravable generarán, de modo que a tasas de tributación suficientemente elevadas será negativa la transferencia neta hacia el resto de la población.[46]

La reflexión sobre la búsqueda de rentas ayuda también a explicar y justificar las limitaciones de la democracia, que se crean por los requerimientos constitucionales y de otra clase de supermayoría o mayoría calificada para la acción legislativa, que discutiremos más adelante. Sin tales limitaciones, la obtención del control de la mayoría de la legislatura sería mucho más valiosa para el ganador, y así las facciones gastarían mayores recursos en la lucha por ese control. Los recursos adicionales se desperdiciarían en gran medida desde un punto de vista social, sobre todo si la meta de los contendientes fuera el logro del máximo poder mediante la extinción de la democracia.

¿Sirve la regulación de las contribuciones y tácticas de las campañas al propósito de convertir a la democracia en un método más o menos eficiente para organizar el mercado político? Promulgada en 2002, la Ley de Reforma de las Campañas Bipartidistas (BCRA, y conocida también como McCain-Feingold) requiere de quienes contribuyen a las campañas, entre otras cosas, que las personas ubicadas más allá de un umbral monetario revelen su identidad ante la Comisión Electoral Federal, donde la información estará disponible al público. Asimismo, la Ley exige que los anuncios publicitarios auspiciados por el candidato mismo o por su comité de campaña reconozcan la existencia de ese patrocinio, de modo que el lector o el espectador no crean que se trate

[46] Casey B. Mulligan, "Economic Limits on 'Rational' Democratic Redistribution", U. Chicago, Graduate School of Public Policy Studies, Working Paper Series 01.7, marzo de 2001.

de anuncios desinteresados. La ley también: *1)* eleva el límite impuesto a las contribuciones de campaña realizadas por determinados individuos (el "dinero de campaña") de 1 000 a 2 000 dólares por candidato y *2)* prohíbe las "contribuciones indirectas", esto es, aquellas dirigidas no a un candidato o a su partido sino a grupos "independientes" que compran publicidad durante las campañas políticas pero no apoyan expresamente a los candidatos oficiales. La mayor parte de esos anuncios están en realidad coordinados estrechamente con las campañas de candidatos particulares.[47]

El propósito de la Ley no consiste en limitar los gastos totales de campaña sino en reducir la influencia de los donadores más acaudalados y, asimismo, lograr que la publicidad de campaña sea menos tendenciosa obligando a que se revele el nombre de sus patrocinadores. Esto es cierto, si bien los anuncios en cuestión poseen ciertos rasgos que los hermanan con la carrera armamentista; en la medida en que la promoción de los productos en competencia está encaminada a contrarrestar mutuamente los efectos de los anuncios de cada uno y por ello no motiva a los consumidores para virar de un producto a otro, no tiene ninguna productividad social neta. No obstante, limitar la publicidad significaría negar a los vendedores el uso pleno de una herramienta competitiva no carente de importancia. Según se señaló antes en XX.3, tendría un efecto particularmente nocivo sobre la nueva entrada a los mercados —ya que un elemento de ingreso reciente podría tener que hacer gastos muy fuertes en publicidad para triunfar sobre las preferencias ya afincadas de los consumidores por marcas establecidas—, como la entrada de nuevos candidatos a los mercados políticos. La Ley McCain-Feingold no limita directamente los gastos del candidato, sino que lo hace indirectamente obligándolo a recaudar dinero de muchos individuos distintos. La limitación fijada sobre la publicidad de campaña, llevada a cabo ya sea de forma directa o indirecta, también opera como un subsidio a los periódicos y otros medios noticiosos. Entre menos publicidad política exista, más dependerá el público de los medios para obtener información y opiniones acerca de las campañas políticas. Esto puede explicar por qué los medios apoyan vigorosamente las reformas del financiamiento de las campañas.

El temor que dio inspiración a las leyes para limitar las contribuciones a las campañas se funda en tres aspectos. Primero, el incentivo para involucrarse en una publicidad fraudulenta y en otros renglones desorientadora es más grande en el mercado político debido a que los votantes se muestran menos cautelosos que los consumidores en la medida en que, a fin de cuentas, para cada individuo es muy poco lo que entra en juego al depositar en la urna un voto correcto. Quizá, por consiguiente, sería mejor que los votantes obtuvieran una mayor parte de la información de los medios a que la obtuvieran

[47] En el caso Citizens United, discutido antes en XXIX.1, la Corte Suprema declaró inconstitucional una disposición importante de la Ley.

de la publicidad política. Pero esto ignora el hecho de que el interés del votante y la participación electoral probablemente aumentarán entre más publicidad de campaña se dé; en la proporción en que la publicidad organizada por un contendiente no se aboque a contrarrestar los efectos del otro, incrementará la demanda de su producto; es decir, incrementará la cantidad de votos.

Segundo, está difundido el miedo a que las contribuciones ilimitadas socaven la elección democrática. Un puñado de personas muy acaudaladas podrían, si se les permitiera hacerlo, suministrarle a un candidato poco atractivo a ojos del conjunto de la población un fondo de reserva bélica integrado por cientos de millones de dólares y, mediante un grueso volumen de publicidad, generar un apoyo electoral formidable a su persona. Los individuos acaudalados tienen la puerta abierta para gastar tantos billetes como deseen en sus propias campañas, lo que en efecto ha dado pábulo a algunas candidaturas bastante excéntricas, y las perspectivas de "comprar" elecciones se ampliarían si esos sujetos pudieran aportar cantidades ilimitadas a las candidaturas de otras personas. Pero lo que esto revela es cuán ineficiente resulta permitirles a los ricos que gasten su dinero sólo en sus propias candidaturas. Así se propicia la subespecialización. Si una persona acaudalada se gana el permiso para dar una aportación a otro candidato, aquello que gaste en la campaña será en promedio una inversión más productiva, toda vez que la riqueza y el don para ser un candidato político efectivamente capaz no van de la mano. Igualmente, forzar a un candidato a recaudar dinero de una multitud de personas en lugar de acudir sólo a individuos ricos limita el campo de selección de los candidatos al hacerlos gastar una cantidad ingente de tiempo en la recaudación de fondos. Al obligarlos a trabajar más duro a fin de reunir la misma cantidad de dinero, de hecho, incrementa el costo de la recaudación de fondos, más allá de que, no obstante, esto se contrarreste posiblemente con el resultado de reunir en promedio ligeramente menos fondos de los que se podrían recabar de otra manera; un aumento en el costo de una mercancía se traducirá en una menor demanda de ella.

Tercero, existe la preocupación de que las contribuciones a una campaña, cuando son cuantiosas, constituyan una forma de soborno apenas velado: las corporaciones y los individuos provistos de una gran fortuna no le habían de dar una cantidad sustancial de dinero a un candidato político sin tener la expectativa de un *quid pro quo*. El soborno es ilegal, pero no lo es la clase de soborno "suave" que consiste en contribuciones de campaña que, como lo sabe el receptor, disminuirán o se extinguirán en el futuro si él vota constantemente en contra de medidas a las que sus contribuidores asignan importancia. Aun así, como lo tenemos claro, una debilidad de la democracia como método de asignación de recursos es que los votos no se ponderan por la intensidad de la preferencia, como ocurre en el mercado, y las contribuciones a una campaña constituyen un método para expresar la intensidad de una preferencia. El problema reside en que los beneficios privados y sociales podrían no coin-

cidir, y entre más concentrados estén los beneficios privados y más difusos sean los beneficios sociales, menos probable será que una asignación de recursos guiada por el peso de las contribuciones a las campañas produzca el óptimo social.

No obstante, en virtud de que los votantes tienen muy pocos incentivos para revisar cuidadosamente los anuncios políticos, es digna de mérito la jugada de tratar de sembrar el escepticismo entre los votantes exigiendo, como lo hace la Ley McCain-Feingold, que se divulguen los nombres de los patrocinadores. Pues entonces podría transparentárseles incluso a los votantes desapercibidos que los patrocinadores se lanzan en pos de una transferencia privada y finalmente ineficiente que ha de ser el fruto de prestar su apoyo a un candidato en particular con sus contribuciones. Otra limitación tangible sobre el funcionamiento de un mercado político completamente libre es la prohibición de la compra de votos. Dado que en unas elecciones generales un solo voto depositado en la urna no brinda esencialmente ninguna oportunidad de modificar los resultados, tiene un escaso valor instrumental. Por lo tanto, muchas personas estarían dispuestas a vender sus votos por sumas triviales; en especial se trata de personas que son elegibles para votar, pero en cuyos planes no entra el de molestarse en hacerlo. Y especialmente si pensaran que otras personas estarían bien dispuestas a vender sus propios votos a un precio módico. Uno puede incluso imaginarse el equivalente político de una oferta en dos niveles (véase XV.9 antes): una oferta de comprar el primer millón de votos a 25 dólares, el siguiente millón a 20 dólares, y así en adelante.

Sin duda alguna, el precio de mercado de un voto sería más elevado entre más grande fuera la demanda de votos por parte de los partidos políticos. Pero, incluso si llegara a una cantidad tan alta como 100 dólares, muchos candidatos y sus partidarios tendrían recursos para comprar millones de votos. Tales compras podrían abrir una profunda brecha entre la opinión pública y los resultados electorales.

XX.8. Democracia, jueces electos y las limitantes del mandato judicial

Si la democracia presenta propiedades de eficiencia significativas y es en todo caso la forma preferida de gobierno, ¿no implica esto que los jueces deberían ser electos? De hecho, tal como lo sabemos, en la mayoría de los estados de nuestra nación los jueces reciben el cargo por una elección.

Un juez electo está sujeto a restricciones que sólo tienen una contraparte débil en otros tipos de magistraturas. La primera y más obvia es que, a condición de que haya sido elegido sólo para un periodo limitado y, por consiguiente, deba lanzarse a la relección si desea continuar en el puesto, está sujeto, como no lo está un juez general provisto de titularidad, a una suerte de revisión de su desempeño. La segunda, y estrechamente vinculada a la ante-

rior, es que el juez electo tiene que mostrarse más sensible a la opinión pública que un juez cuya titularidad no depende de los caprichos del electorado. Sólo un puñado de casos, fundamentalmente aquellos que involucran delitos de acusada notoriedad, suscitarán el interés de una porción significativa del electorado, pero en tales casos podemos nuevamente esperar que se disemine de modo gradual un sesgo sistemático. Por ejemplo, en un estado que cuenta con la pena capital, en virtud de que solamente los asesinos más notorios son elegibles para ese castigo, los jueces pueden inclinarse en contra de los demandados por asesinato.[48] Adicionalmente, podemos esperar que los jueces electos se inclinen en mayor proporción que los jueces designados a favor de un abogado litigante que sea residente del estado del juez cuando la parte contraria es encabezada por un no residente.[49]

Tercera, el juez tiene que demostrar que posee la habilidad para recaudar dinero a fin de llevar adelante su campaña electoral, y los donadores fundamentales de las campañas para la elección de magistrados son los abogados que litigan en el tribunal del juez. Si los abogados situados a ambos lados de las áreas de práctica principales —tales como los que están al pendiente de casos de pacientes víctimas de una mala práctica médica y los que están al pendiente de casos de demandados por una mala práctica médica— dieran una cantidad equitativa de dinero a los candidatos judiciales, la situación sería muy similar a la que atañe a los árbitros: el juez tendría un incentivo para contemporizar con ambas partes y hallar una solución intermedia al emitir fallos sobre esos casos. Con todo, en las áreas de práctica particulares sucede con frecuencia que los intereses en juego son sistemáticamente asimétricos, y en un escenario de esa índole es probable que el magistrado electo presente un sesgo marcado.

Si atamos los hilos de este asunto, podemos esperar que los jueces que han sido elegidos sean políticamente menos independientes que los que han sido nombrados, en especial los jueces por designación que disponen de una titularidad de por vida. Aun así, la condición de los primeros no es por necesidad una mala cosa, y ello no sólo debido al acicate para esforzarse que les supone la falta de seguridad en el empleo. Por añadidura, las decisiones de los jueces electos tienden a ser más predecibles que las de los designados.[50] La razón estriba precisamente en que los jueces electos tienen menos independencia; el juez independiente probablemente tendrá que hacer un cálculo más com-

[48] En consonancia con esta conjetura, entre más breve sea el periodo promedio de permanencia de los jueces en un estado, más probable será que impongan la pena de muerte. Paul R. Brace, "The Interplay of Preferences, Case Facts, Context, and Rules in the Politics of Judicial Choice", 59 *J. Pol.,* 1206 1221, 1223 (1997).

[49] Para evidencias, véase Alexander Tabarrok y Eric Helland, "Court Politics: The Political Economy of Tort Awards", 42 *J. Law & Econ.,* 157 (1999).

[50] F. Andrew Hanssen, "The Effect of Judicial Institutions on Uncertainty and the Rate of Litigation: The Election *versus* Appointment of State Judges", 28 *J. Legal Stud.,* 205 (1999).

plejo a la hora de tomar decisiones, puesto que no puede simplemente pararse a ver por dónde soplan los vientos políticos. Más aún, si la existencia de una magistratura electiva significa que hay una preferencia democrática legítima por crear una mayor concordancia entre las actitudes judiciales y las populares que en un sistema no electivo, entonces no sólo será menos probable que se gane la relección un juez que desafía a la opinión pública, sino que, paradójicamente, será (cabe sostenerlo) un mal juez, incluso un usurpador. El otro lado de la moneda, empero, es que, entre más uniforme es la opinión pública, más importante resulta la independencia judicial en lo concerniente a la salvaguarda de los derechos de las minorías. Es probable que esta independencia sea muy valorada por los políticos allí donde la competencia política es intensa debido a que, "al establecer un tribunal independiente, los políticos que aún están en funciones les ponen más barreras a los sucesores afanosos de alterar las políticas aprobadas el día de hoy".[51]

No deja de ser curioso que, aunque una magistratura electiva sea más democrática que la establecida por designación en el ámbito angloestadunidense, no sería más democrática que una magistratura de carrera, tal como la que se halla en los sistemas legales que no son de orígenes angloestadunidenses. Cuando los códigos legislativos son amplios y los jueces son formalistas en el sentido de que los hacen cumplir tal como están escritos en vez de usarlos como el punto de partida para el desarrollo de estándares legales, la legislatura democrática es la que lleva la batuta legal y en realidad los jueces simplemente están ejecutando las decisiones tomadas en el curso del proceso democrático.

La magistratura electiva limita el campo de selección. Por motivos de temperamento, la mayoría de las personas no cuentan con las aptitudes para navegar en el mar de la política electoral y, en todo caso, no tienen la destreza requerida, aunque pueden reunir las habilidades requeridas para ser un juez de excelencia. Probablemente sea muy reducido el número de personas que tienen a la vez talento político y talento judicial, y aun puede ser que se dé un grado de incompatibilidad entre los dos tipos de talento.

Los jueces federales no sólo no son elegidos: son designados de por vida, y resulta extremadamente difícil removerlos. Esto les presta un aire monárquico; y existe una oposición fundada en sólidos principios al grado en que la Corte Suprema de los Estados Unidos ha campeado por sus fueros con todo su peso al tratar asuntos políticos sumamente controvertidos. ¿No podría hallarse una respuesta en la imposición de límites a un periodo? Y es que, ya sea en el ámbito de la academia o el de la magistratura, una titularidad de por vida es una fórmula óptima para el abuso. Elimina cualquier castigo por escurrirle el bulto al trabajo; y la estructura salarial de los jueces federales (a

[51] F. Andrew Hanssen, "Is There a Politically Optimal Level of Judicial Independence?", 94 *Am. Econ. Rev*., 712 (2004). Véase también J. Mark Ramseyer, "The Puzzling (In)Dependence of Courts: A Comparative Approach", 23 *J. Legal Stud.,* 721 (1994).

todos los jueces de la misma jerarquía se les paga lo mismo, y existen muy pocos rangos) proscribe el uso de la zanahoria junto con el garrote. Esto sugiere que posiblemente sea de desear que se fijen límites a los periodos de los jueces: digamos, 10 años (que es un plazo común para los jueces de tribunales constitucionales en otros países). De este modo se acortaría la duración del servicio de los afectos a gandulear y también se crearía un incentivo para el buen desempeño, ya que el juez anhelaría asegurarse de conseguir un buen trabajo después de que expirara su periodo. El lado negativo de esas disposiciones es que los magistrados llegarían a distraerse de sus faenas al tener que hacer arreglos para obtener otro empleo al final de su periodo; sus decisiones podrían salir distorsionadas a causa del deseo de conquistar el favor de los empleadores potenciales en el futuro a base de una actitud obsequiosa; una rotación más acelerada de los jueces reduciría la estabilidad legal; y resultaría más difícil atraer a candidatos capaces al reducir la flexibilidad de su carrera. Empero, estos argumentos a favor de fijar límites podrían ser contundentes si no fuera por el gran rigor y cuidado observado en el proceso de selección de los candidatos a la magistratura federal, el cual elimina a aquellos más proclives a volverse gandules, en especial dado que la mayoría de esos jueces son elegidos cuando alcanzan la edad mediana y ya existe una información abundante acerca de su ética de trabajo.

En cambio, abogar a favor de imponer límites al periodo de duración de los jueces de la Corte Suprema tiene bases más sólidas. La Corte Suprema es en una gran proporción, si no es que enteramente, un tribunal político —integra casi una tercera rama de la legislatura— y la titularidad otorgada de por vida a los políticos es profundamente antidemocrática. Además, en vista de que ha aumentado su longevidad, es probable que los jueces presten servicio durante periodos muy prolongados hasta llegar a una edad muy avanzada; esto fortalece el argumento en cuestión.

XX.9. Un enfoque antimonopólico de la democracia

La teoría económica de la democracia apunta hacia los posibles procedimientos para limitar los malos efectos de la estructura duopólica del sistema político estadunidense. El peligro de colusión entre los competidores, incluidos los partidos políticos rivales, es grave cuando hay sólo dos, a menos que la colusión provoque la formación y el rápido crecimiento de un tercer partido, de modo que se corrigiera automáticamente. La calidad y la sensibilidad de las políticas y los candidatos que ofrecen al electorado dos partidos ideológicamente similares tenderán a alejarse del óptimo si no hay ninguna amenaza de entrada creíble de un tercer partido que pueda ofrecer mejores políticas y candidatos. La importancia de los terceros partidos reside así en que mantienen suspendidos de un hilo a los partidos principales, y aun, en una analogía

con la "competencia por el mercado" (económico) (por ejemplo, la competencia para obtener una franquicia exclusiva de televisión de cable que conferirá un poder monopólico al concesionario; véase antes XIV.11), en que aseguran la eliminación ocasional de uno de los partidos principales y su sustitución por un partido minoritario. Aunque eso no ocurra, uno o ambos de los partidos principales podrían copiar las políticas de un tercer partido e incluso "quitarles" a sus candidatos, lo que es algo bueno.

Sin embargo, entre ellos los dos partidos controlan el sistema político de todos los estados de la Unión Americana y pueden promulgar cualesquiera leyes en su interés conjunto. Por más vigorosamente que compitan entre sí para ganar las elecciones, tienen un interés común en cerrar el paso a la entrada de un tercer partido, a menos que uno de los partidos principales confíe en que el tercero le quitará más votos al otro partido principal que a él mismo.

No es que la entrada de un tercer partido a la arena electoral deba dejar de ser costosa. Si no hay *ciertas* barreras para terceros partidos, las listas electorales contendrían tantos candidatos que los votantes se verían confundidos. La confusión es una objeción contra el permiso para que todos los candidatos participen en todos los debates televisados.

También debieran verse con una mirada crítica las medidas que protegen a los funcionarios actuales por oposición a la protección de los partidos principales, limitando así la competencia electoral incluso dentro del campo reducido que permite un sistema bipartidista. Los funcionarios actuales tienen fuertes ventajas naturales en esa competencia, semejantes a las de los monopolistas arraigados en los mercados económicos. La dificultad para que haya una nueva entrada es particularmente grave en los mercados políticos en virtud del carácter de "el ganador lo toma todo" que posee la competencia política. Un nuevo miembro en un mercado económico que obtiene una participación de 10% podría sobrevivir muy bien; un candidato que obtiene 10% de la votación o, para el caso (si sólo hay dos candidatos), 49% es un perdedor, y punto.

Aun cuando la limitación de los periodos para los legisladores aminora las ventajas a favor de quienes ocupan el puesto en la actualidad, también tiene su lado negativo. Entre menor sea el periodo esperado de un legislador en el cargo, menos probable será que apoye la legislación que implique beneficios futuros pero costos presentes. Las limitaciones de los periodos también reducen el incentivo para adquirir habilidades políticas al disminuir la ganancia derivada de su adquisición, y crean más funcionarios debilitados a causa de su retiro inminente, que probablemente serán insensibles a los intereses del electorado (un problema del "último periodo"). Pero el problema del funcionario o estadista debilitado no debe exagerarse. Si fuera en verdad grave, ¿por qué relegiría el electorado a alguien (el presidente, por ejemplo) que fuera inelegible para postularse al término de su nuevo periodo?[52] La

[52] Linda Cohen y Matthew Spitzer, "Term Limits", 80 *Geo. L. J.*, 477 (1992).

respuesta podría ser que el candidato no habría sido elegido la primera vez si sus políticas no hubiesen estado de acuerdo con las de sus electores, y, de ser así, puede esperarse que ese acuerdo continúe, aunque no esté planeando presentarse para la relección.

Los referendos y, en particular, las iniciativas, que son los instrumentos principales de la democracia directa en los gobiernos modernos, sirven al propósito, importante en un sistema bipartidista, de permitir el rompimiento de los tratos colusorios entre los partidos.[53] Esta observación nos recuerda el carácter antimonopólico de las políticas que estamos discutiendo. La iniciativa es un método para la "integración hacia atrás", por el que los clientes destruyen a los cárteles de proveedores entrando al mercado cartelizado y compitiendo con los proveedores existentes.

La política antimonopólica asume formas diferentes; ¿cuál de ellas es más apropiada para la regulación del proceso democrático? Comparemos dos modelos. Uno, que es estático, destaca el efecto benéfico sobre la asignación de los recursos de tener cierto número de empresas rivales. El efecto principal destacado es el de hacer bajar el precio hasta el nivel del costo marginal (véase antes IX.3). La analogía en la arena política sería un sistema en el que hay cierto número de partidos políticos, cada uno de los cuales representa los intereses de cierto segmento de la población, y que en conjunto representan a todos esos intereses y maximizan el bienestar (en el sentido de la contemporización con los intereses), al menor costo posible, mediante una solución intermedia alcanzada entre las partes.

Un modelo dinámico contrastante se origina en el famoso concepto de Schumpeter de "el huracán de la destrucción creativa" (véase antes IX.4). Como recordaremos, Schumpeter sostuvo que el bienestar económico se maximiza a lo largo del tiempo como resultado de una sucesión de monopolios. Lo importante de su análisis no es el número de los competidores que haya en cualquier momento dado, sino que los competidores existentes no pueden atrincherarse contra las nuevas entradas. Y es muy posible que eso sea lo más importante en los mercados políticos: no que haya muchos partidos sino que los partidos nuevos (como el Partido Republicano de Lincoln, en 1860), o las coaliciones nuevas dentro de los partidos existentes (como las coaliciones que dieron a Franklin Roosevelt sus victorias electorales presidenciales), no se vean obstruidos por las formaciones políticas existentes.

[53] Véase Bruno S. Frey, "Direct Democracy: Politico-Economic Lessons from Swiss Experience", 84 *Am. Econ. Rev. Papers & Proceedings,* 338 (mayo de 1994). Técnicamente, un referendo es un voto popular sobre una propuesta que ya ha sido aprobada por la legislatura; una iniciativa es un voto popular que se realiza sobre una propuesta y que es sometido directamente a la aprobación de los votantes sin mediación legislativa.

Lecturas recomendadas

Anderson, Gary M. y Robert D. Tollison, "Congressional Influence and Patterns of New Deal Spending, 1933-1939", 34 *J. Law & Econ.,* 161 (1991).
Becker, Gary S., "Pressure Groups and Political Behavior", en *Capitalism and Democracy: Schumpeter Revisited,* 120 (R. D. Coe y C. K. Wilbur [comps.], 1985).
Buchanan, James M. y Gordon Tullock, *The Calculus of Consent: Logical Foundations of Constitutional Democracy* (1962).
Cooter, Robert D., "The Objectives of Public and Private Judges", 41 *Pub. Choice,* 107 (1983).
Downs, Anthony, "In Defense of Majority Voting", 69 *J. Pol. Econ.,* 192 (1961).
Epstein, Lee, William M. Landes y Richard A. Posner, *The Behavior of Federal Judges: A Theoretical and Empirical Study* (2013).
Garrett, Elizabeth, "Term Limitations and the Myth of the Citizen-Legislator", 81 *Cornell L. Rev.,* 623 (1996).
Kalt, Joseph P. y Mark A. Zupan, "The Apparent Ideological Behavior of Legislators: Testing for Principal-Agent Slack in Political Institutions", 33 *J. Law & Econ.,* 103 (1990).
Landes, William M. y Richard A. Posner, "The Independent Judiciary in an Interest-Group Perspective", 18 *J. Law & Econ.,* 875 (1975).
Lim, Claire S. H., "Preferences and Incentives of Appointed and Elected Public Officials: Evidence from State Trial Court Judges", 103 *Am. Econ. Rev.,* 1360 (2013).
McChesney, Fred S., "Rent Extraction and Rent Creation in the Economic Theory of Regulation", 16 *J. Legal Stud.,* 101 (1987).
McNollgast (Matthew McCubbins, Roger Noll y Barry Weingast), "Politics and the Courts: A Positive Theory of Judicial Doctrine and the Rule of Law", 68 *S. Cal. L. Rev.,* 1631 (1995).
Niblett, Anthony, Richard A. Posner y Andre Shleifer, "The Evolution of a Legal Rule", 39 *J. Legal Stud.,* 325 (2010).
Olson, Mancur, Jr., *The Logic of Collective Action: Public Goods and the Theory of Groups* (1965).
Posner, Richard A., "Theories of Economic Regulation", 5 *Bell J. Econ. & Mgmt. Sci.,* 335 (1974).
——— *How Judges Think* (2008).
——— *Reflections on Judging* (2013).
Ramseyer, J. Mark, "The Puzzling (In)Dependence of Courts: A Comparative Approach", 23 *J. Legal Stud.,* 721 (1994).
Research Handbook on Public Choice and Public Law (Daniel A. Farber y Anne Joseph O'Connell [comps.], 2010).

Rubin, Paul H., "Micro and Macro Legal Efficiency: Supply and Demand", 13 *S. Ct. Rev.,* 19 (2005).
Stevenson, Matthew C., "Legal Realism for Economists", *J. Econ. Perspectives,* primavera de 2009, 191.

Problemas

1. ¿Debiera prohibirse que un juez federal actuara como uno de los directores de un fondo mutuo diversificado? ¿Que fuera propietario de acciones del fondo? Si en efecto es propietario de tales acciones, ¿debiera excusarse de participar en juicios que impliquen a compañías representadas en la cartera del fondo?

2. ¿Cómo se alteraría el comportamiento judicial si los sueldos de los jueces fueran un porcentaje de los honorarios de los abogados litigantes por la presentación de los juicios?

3. Si los salarios en el ámbito judicial fueran mucho más elevados de lo que ya son, ¿cómo afectaría ese incremento a la selección de los jueces?

4. ¿Debiera permitirse que el Congreso aumentara los sueldos de los jueces durante sus periodos vigentes?

XXI. EL PROCESO DE ELABORACIÓN DE LAS NORMAS JURÍDICAS

XXI.1. El cuerpo de precedentes como un acervo de capital

Este capítulo continúa la discusión comenzada en el anterior sobre los procesos mediante los que se crea la ley y sobre las instituciones que la crean. Se hace hincapié en el papel de los jueces. Son ellos los autores del derecho común; no obstante, en el sistema legal estadunidense, incluso en los campos estatutarios, muchas de las normas de obligación legal son glosas judiciales a partir de las promulgaciones realizadas por la legislatura.

Una decisión tomada por un tribunal de apelaciones establece un precedente,[1] una base o causa para decidir respecto de un caso en la misma forma. Con todo, la eficacia del precedente depende —por paradójico que parezca— del poder que tengan los jueces para anular un precedente. Si no puede ser anulado, el hecho de que los casos subsecuentes se amolden a él no es evidencia de que los jueces involucrados en ese caso lo consideren sólido, y, de esta forma, la acumulación de precedentes que atañen a la misma cuestión no transmitirá información alguna acerca de la solidez de la norma que establezcan los casos; mientras que, si los jueces hubieran podido anular el precedente original pero no lo hicieron, ésta es una prueba de que lo consideran sólido.[2] Pero sería más preciso decir que esto confirma meramente que consideraban al precedente lo bastante sólido para no anularlo. Podrían preferir que el caso se hubiera decidido de manera diferente, pero el valor economizador de la decisión fundada en el precedente se perdería si los jueces reconsideraran los precedentes en cada uno de los casos.

El cuerpo de precedentes en un área de la ley puede ser concebido como un acervo de bienes de capital; específicamente, un acervo de conocimientos que rinde servicios durante muchos años a litigantes potenciales bajo la guisa de información acerca de obligaciones legales. Los bienes de capital se deprecian; el valor de los servicios que rinden declina a lo largo del tiempo. Esto puede deberse al desgaste físico del bien o a la obsolescencia: un cambio en el ambiente que reduce el valor de los servicios que genera el bien. El primer tipo de depreciación carece de importancia en el caso de la información, pero el segundo es muy importante; el derecho de los accidentes que se desarrolló

[1] A menos que (en el sistema estadunidense) el tribunal juzgue que la decisión no tiene carácter de precedente.

[2] Todd J. Zywicki, "The Rise and Fall of Efficiency in the Common Law: A Supply-Side Analysis", 97 *Nw. U. L. Rev.,* 1551, 1577 (2003).

para afrontar las colisiones entre los carros tirados por caballos será menos valioso cuando se aplique a las colisiones de automóviles.

Por supuesto, el hecho de que un bien de capital se deprecie no significa que el *acervo* de capital del que forma parte deberá contraerse. Ello dependerá de la tasa de remplazo de los bienes de capital. A medida que los viejos precedentes se tornan obsoletos y finalmente dejan de ser parte de la existencia de precedentes útiles, por medio de los litigios se añaden nuevos precedentes a ese acervo.

Las tasas de depreciación de los precedentes pueden calcularse usando el número de veces que se citen al tomar una decisión como sustituta del valor de la decisión como precedente.[3] Resulta que los precedentes se deprecian más rápidamente en un campo en el que hay mucha actividad de creación de leyes estatutarias —un precedente que se basa en la interpretación de las leyes estatutarias probablemente se volverá obsoleto cuando ese lenguaje cambie—, y el capital jurídico general se deprecia más lentamente que el capital jurídico específico, porque cuanto más general sea una regla, tanto menos probable es que un cambio técnico o jurídico la vuelva obsoleta; la decisión en el caso de Hadley *vs.* Baxendale[4] es un buen ejemplo. Los precedentes de la Corte Suprema de los Estados Unidos se deprecian más lentamente que los de los tribunales federales de apelaciones, porque la Corte Suprema es más selectiva en su elección de los casos que revisará que los tribunales de apelaciones (en realidad, a diferencia de la Corte Suprema, un tribunal federal de apelaciones no puede negarse a decidir sobre un caso que caiga dentro de su jurisdicción), y es aceptable suponer que la Corte Suprema ejercerá su selección en favor de los casos más generales a fin de que le rinda más el delgado fajo de billetes que le dan. La teoría alternativa de que los precedentes de la Corte Suprema se deprecian más lentamente porque tienen más autoridad, es decir, son más valiosos, no es aceptable económicamente. La tasa a la que se deprecia un bien no es una función de su valor. Las computadoras se deprecian más rápidamente que los desatornilladores.

XXI.2. La producción de precedentes

El acervo de capital de precedentes es el producto conjunto de los abogados y los jueces que participan en la argumentación y la decisión, respectivamente, de los juicios, en particular en los casos de apelación. Una característica extraña de este proceso de producción es que no se paga a los productores. Ni los jueces ni los abogados de Hadley *vs.* Baxendale recibieron alguna regalía u otra compensación por un precedente que ha guiado la decisión de

[3] William M. Landes y Richard A. Posner, "Legal Precedent: A Theoretical and Empirical Analysis", 19 *J. Law & Econ.,* 249 (1976).

[4] 9 Ex. 341, 156 Eng. Rep., 144 (1854). Véase antes IV.11.

varios miles de juicios y, lo que es aún más importante, ha estructurado las relaciones comerciales de millones de compradores y vendedores. Pero, antes de concluir que la producción de precedentes es subóptima, debemos considerar el carácter de los precedentes como un subproducto del proceso de litigación. Gran parte del humo fabril se genera aun cuando no haya ningún mercado para él, gracias a la demanda de bienes cuya producción genera el humo como un subproducto. Lo mismo ocurre con los precedentes y la litigación.

El arreglo extrajudicial es más barato que la litigación. Por lo tanto, sólo si cada uno de los litigantes espera ganar en el litigio más de lo que el otro espera que gane, es probable que las partes no puedan llegar a un arreglo que haga a ambas sentirse en mejor situación de la que esperaban lograr en el litigio (véase después en XXII.5). La incertidumbre es una condición necesaria de tal divergencia de cálculos. Puede ser fáctica o legal, pero sólo la incertidumbre legal es relevante aquí. Si es grande, habrá mucha litigación, incluida mucha apelación. Pero dado que la litigación, en particular al nivel de apelación, genera precedentes, su aumento conducirá a una reducción de la incertidumbre jurídica, haciendo que la cantidad de la litigación baje en el periodo siguiente. Finalmente, cuando se producen pocos precedentes nuevos, la incertidumbre aumentará a medida que se deprecian los precedentes antiguos (porque son menos informativos en un ambiente cambiado), y esta incertidumbre suscitará un nuevo aumento de la litigación y, por lo tanto, una producción mayor de precedentes. Así pues, aun cuando no haya mercado para los precedentes como tales, su producción aumentará cuando su valor social aumente a resultas de la mayor incertidumbre y disminuirá a medida que ese valor baje.

Este análisis supone que los jueces están interesados en la producción de precedentes; tal supuesto es una extensión natural del modelo del comportamiento judicial esbozado en el capítulo anterior. Decidir un caso aplicando un precedente economiza el tiempo y el esfuerzo del juez; no tiene que replantearse el problema inherente al caso partiendo de los cimientos. Además, un precedente proyecta la influencia de un juez con mayor eficacia que una decisión que no tendrá ningún efecto en la orientación del comportamiento futuro. Esto sugiere también la razón de que los jueces sigan los precedentes además de hacerlos[5] y, por ende, la razón de que los abogados defiendan casos sobre las bases del precedente. Otro motivo por el que el juez sigue el precedente es que, si la actual generación de jueces no sigue este último, es menos probable que la próxima generación siga los precedentes de la generación actual, porque es menos probable que los jueces de la generación siguiente sean

[5] Para evidencias de que los jueces sí siguen en realidad el precedente incluso cuando la probabilidad de revocación es baja, véase Frank Cross, "Appellate Court Adherence to Precedent", 2 *J. Empirical Legal Stud.*, 369 (2005). Por supuesto, la probabilidad puede ser baja debido a que los jueces quedan disuadidos ante la amenaza de que se dé una revocación por haber pasado por alto el precedente.

criticados por no apegarse a las decisiones de sus predecesores. Son moderados los costos de la crítica profesional para los jueces, pero, dado que las normas de la permanencia con titularidad y la compensación judiciales atenúan los incentivos que operan habitualmente sobre los seres humanos, es probable que los jueces se vean influidos por lo que en la mayoría de las áreas de la vida es una fuerza débil (la crítica).

Es cierto, con todo, que hay un problema de "gorroneo". El juez que desatiende un precedente en su celo por establecer su propio precedente en contrario podría dar al traste sólo trivialmente con la práctica de la decisión basada en el precedente; los costos privados de su acto pueden ser menores que las ganancias privadas. Pero la estructura de la revisión de apelación frena el problema del "gorroneo". El juez que desatiende un precedente será revocado por un tribunal superior, que no tiene ningún interés en permitir que *él* dé al traste con la práctica de la decisión basada en el precedente, a fin de magnificar su propia influencia. En cada jurisdicción hay un tribunal supremo cuyas decisiones no están sujetas a una revisión posterior. Y dentro de cada tribunal disminuye el problema del "gorroneo". Si los magistrados de la Corte Suprema desatienden el precedente en sus decisiones, sabrán que están reduciendo la probabilidad de que los magistrados futuros les otorguen algún peso a sus propias decisiones como un precedente.

Otros factores que impulsan a los jueces a seguir el precedente son que reduce sus costos de decisión, que al restringir la discrecionalidad judicial les resta base a las percepciones de los jueces como legisladores políticamente irresponsables y que pasar por encima del precedente aumenta la cantidad de litigación al disminuir la certeza legal, y el aumento de la litigación generaría presiones para añadir jueces nuevos (diluyendo así la influencia de los jueces existentes), para hacer que cada juez trabajara más arduamente, o para sustituir a las cortes por los tribunales arbitrales y administrativos, lo que disminuirá la influencia del poder judicial.

Los jueces aplican los precedentes y algunas veces los anulan, pero por añadidura optan frecuentemente por "distinguir" precedentes, y —cuando "distinguir" no es, como habitualmente se le considera, un simple eufemismo de "anular"— ésta es una importante técnica para dar más eficiencia a las vías del derecho y la ley.[6] Distinguir un caso equivale a aceptar su principio básico pero calificarlo a la luz de la información obtenida a través de casos subsecuentes. De esta manera, el primer caso en una serie podría fundamentar la decisión de que un gesto amenazador está sujeto a proceso como un asalto civil, pero el segundo caso —uno en el que el demandado le había dicho al demandante: "Si no tuvieras más edad que yo, hundiría este lápiz afilado en tu ojo"— podría calificar el principio al establecer que un gesto amenazador

[6] Véase Nicola Gennaioli y Andrei Shleifer, "The Evolution of Common Law", 115 *J. Pol. Econ.,* 43 (2007).

no está sujeto a proceso legal si la amenaza queda refutada por las palabras que la acompañan. El acto de distinguir preserva la captación profunda de los jueces del caso anterior, pero añade a ello las captaciones profundas que tuvieron al respecto los jueces que intervinieron en casos posteriores sobre la base de los diferentes hechos (la información) examinados en esos últimos casos. En contraste, anular un precedente rebaja su valor a cero, y una práctica de anulación frecuente desestabiliza al derecho.[7]

La función de distinguir al amplificar la base informativa de la ley es una invitación a abrazar la diversidad judicial. Los diversos juegos traen distintas perspectivas a cuento, con lo que crecen las probabilidades de distinguir un precedente y, al proceder así, refinar la ley existente.

XXI.3. Producción estatutaria; normas contra estándares

La producción de estatutos o leyes podría parecer un proceso simple por comparación con la producción de precedentes como un subproducto de la litigación. Pero entonces, ¿por qué han subsistido tantas leyes elaboradas por los jueces y menos por los cuerpos legislativos? De hecho, son elevados los costos de la producción estatutaria de normas. La promulgación de un estatuto o ley requiere el acuerdo de una mayoría de los legisladores, y los costos de transacción son altos cuando hay cientos de partes involucradas en la transacción. Esto significa también que la capacidad productiva de una legislatura resulta difícil de ampliar. En efecto, la adición de legisladores podría reducir esa capacidad al aumentar los costos de transacción de la promulgación. Pero la reducción del número de legisladores, aun cuando podría aumentar el número de los estatutos que la legislatura podría producir al reducir los costos de cada promulgación, aumentaría los costos de agencia. El electorado para cada legislador sería más nutrido y esto aminoraría el incentivo del votante individual para monitorear el desempeño de su representante. Y quizá no deseemos que la acción de legislar sea algo tan fácil (véase después XXV.2).

La capacidad limitada de las legislaturas ayuda a explicar el enorme crecimiento de las agencias administrativas en el siglo pasado (en el capítulo XXIV se ofrece una explicación suplementaria de ello) y también ayuda a explicar por qué los estatutos son a menudo tan ambiguos. Un procedimiento para disminuir el tiempo invertido en la negociación consiste en convenir en menos: dentro del contexto legislativo, en dejar las cuestiones difíciles para la resolución futura por parte de los tribunales. Esto significa, a su vez, que si los tribunales adoptan una política de interpretación estrecha de la legislación (construcción estricta), reducirán la producción efectiva de la legislatura al reducir el al-

[7] Véase Nicola Gennaioli y Andrei Shleifer, "Overruling and the Instability of Law", 35 *J. Comparative Econ.*, 309 (2007).

cance o amplitud de cada estatuto. Así, los jueces que son literalistas interpretativos no son los servidores fieles de la legislatura que pretenden ser. Antes bien, incrementan el costo del proceso legislativo al hacer que los legisladores tengan que legislar para cubrir el mismo terreno. Una paradoja relacionada con esta cuestión es que, entre más costosa es la promulgación legislativa, más probable es que sean más amplias las normas estatutarias y, por consiguiente, mayor será la libertad de interpretación de los jueces.

Los tribunales también tienen que elegir entre la elaboración de normas amplias y estrechas, así que convendrá que consideremos más sistemáticamente que en XIV.1 el intercambio de compensaciones y sacrificios entre las normas amplias (generales) y las estrechas (específicas): por ejemplo, entre prohibir la conducción irrazonablemente rápida y prohibir la conducción en exceso de un límite de velocidad especificado. Las normas tienden a ser simples y claras, lo que reduce los costos de su cumplimiento estricto y les facilita a los observadores la evaluación del tribunal o de otra agencia que las aplica a hechos particulares. La simplicidad de las normas y la facilitación brindada al monitoreo de su cumplimiento prestan un especial atractivo a las sociedades en las que la magistratura es propensa a la incompetencia y la corrupción.[8] Aun así, sucede con frecuencia que las normas no son *realmente* simples y claras. A consecuencia de la presión que se ejerce para dar lugar a excepciones y los problemas de delimitación que éstas crean, puede no resultar claro si un caso particular cae dentro de la norma general o dentro de una de sus excepciones.

Las normas tienden a ser crudas, puesto que excluyen los hechos relevantes (tales como, en el ejemplo del límite de velocidad, las características del tráfico, el clima y la hora del día); de este modo, se apoyan en una base informativa más estrecha que los estándares. Dicha exclusión también hace que de alguna forma sean arbitrarias, y, a resultas de ello, carezcan en apariencia del mínimo fundamento racional. La de "ser cuidadoso" tiene a primera vista justificación lógica; la de manejar a menos de 80 kilómetros por hora no la tiene, razón por la cual es necesario poner señalamientos sobre el límite de velocidad. Las normas, en contraste con los estándares, tienden asimismo a separar la creación de las reglas de su aplicación: las legislaturas promulgan las normas; los tribunales las aplican. Los tribunales de derecho común a un tiempo crean y aplican estándares, y al dejar ambas funciones en manos de la misma organización se obtienen ganancias de eficiencia.

Las desventajas de las normas vienen a ser las ventajas de los estándares. Los estándares son flexibles, tienen a todas luces una justificación racional, y generan y hacen uso de más información, incluida aquella que está disponible sólo después de que el estándar se ha adoptado y aplicado; un problema

[8] Véase Barbara Luppi y Francesco Parisi, "Rules *versus* Standards", en *Production of Legal Rules,* 43 (Francesco Parisi [comp]., 2011).

grave de las normas es que excluyen de toda consideración aquellos factores de cuya relevancia no fue posible percatarse a la hora en que se promulgaron. Los estándares también simplifican la operación de fusionar al hacedor con el aplicador de un estándar: a menudo se trata de la misma entidad, a saber, del mismo tribunal. Ahora bien, también es cierto que los estándares son vagos; que su forma de administrarlos es costosa en la medida en que son de interpretación abierta; y que resulta difícil monitorear el cumplimiento que de ellos hagan el tribunal u otro cuerpo a cargo de su aplicación. Entre más sean los tribunales que inspiren desconfianza, ya sea porque haya sospechas de corrupción, de incompetencia o de carencia de recursos para determinar los hechos de manera exacta, más atractivo cobrarán las normas como fuentes de la ley a ser aplicada por los tribunales.

El control de la conducta mediante un conjunto de normas detalladas y no mediante un estándar general supone costos tanto por traducir el estándar de un conjunto de normas legales a un conjunto de normas precisas como por hacer la revisión periódica para mantenerlas al tanto de las condiciones cambiantes; como hemos señalado, una norma específica se volverá obsoleta con mayor rapidez que un estándar general. Los costos de la gobernanza mediante normas específicas son particularmente elevados para organismos como la Corte Suprema o el Congreso, donde es muy costosa la promulgación de cada regla. Pero los beneficios de la particularización superan a menudo los costos. Estos beneficios se obtienen en tres niveles: en la guía de los tribunales mismos, en la guía del comportamiento de las personas sujetas a la norma y en la guía del comportamiento de las partes involucradas en las disputas reales. (Pospondremos la discusión del último de estos beneficios hasta el próximo capítulo.)

1. Podemos imaginar que el derecho común tenga un solo estándar de decisión —alguna versión de la Fórmula de Hand— que los tribunales aplicarían de nuevo en cada caso. Pero serían enormes los costos de información de los jueces y jurados al determinar el resultado eficiente de cada disputa. La existencia de normas específicas limita el alcance y, por ende, el costo de la investigación judicial; en términos económicos, la búsqueda de un máximo local sustituye a la búsqueda más escurridiza de un máximo global.
2. Si una ley es poco clara, los infractores potenciales descontarán (multiplicarán) el costo del castigo no sólo por la probabilidad de que sean atrapados (véase antes VII.2), sino también por la probabilidad adicional, considerablemente menor a uno, de que se determine que la ley será aplicable a la conducta observada por ellos. En consecuencia, disminuirá el efecto disuasivo de la ley. De igual modo, la vaguedad de la ley creará un riesgo de que se determine que la conducta legítima viola la ley. Esto disminuirá más aún el efecto disuasivo, porque tal

efecto depende de la diferencia existente entre el costo del castigo esperado de la conducta ilegal y de la conducta legal.[9] También impedirá algunas actividades legítimas, como destacamos en el capítulo 7; así, el punto es no que la vaguedad reduzca el efecto disuasivo total de la ley, sino que reduce el efecto disuasivo *intentado*. El efecto disuasivo no intentado es un efecto particularmente significativo cuando la actividad legítima disuadida por la prohibición vaga es más valiosa desde el punto de vista social que del privado. Entonces, incluso un aumento ligero de sus costos (debido a la amenaza del castigo) podría reducir grandemente la demanda privada de la actividad, lo que podría imponer sustanciales costos sociales. Esto podría explicar el hecho de que la Corte Suprema haya sostenido que los estatutos penales demasiado amplios son constitucionalmente inválidos cuando la conducta regulada por el estatuto se relaciona estrechamente con la expresión de ideas.[10] El valor social de las ideas supera a menudo su valor privado (como veremos en el capítulo XXIX), dada la ausencia de un sistema adecuado de derechos de propiedad de las ideas.

Este análisis sugiere también una razón de que en los estatutos penales se tolere menos vaguedad que en los estatutos civiles, aun cuando no haya ningún peligro para la libertad de expresión. Como vimos en el capítulo VII, las sanciones penales son más costosas que las civiles, de modo que son mayores los costos sociales del castigo de la conducta legítima a resultas de la vaguedad de la regla legal, lo que justifica una inversión mayor en la particularización de la ley. Es también relevante aquí que no es posible asegurarse contra las sanciones penales. Incluso una probabilidad pequeña de castigar una actividad legal podría inducir a quienes sientan aversión por el riesgo a desechar la actividad, aunque su valor social pudiera ser sustancial.[11]

Entre más exacta y detallada sea una norma, más probable será que contenga fisuras o lagunas: que se permita por implicación la conducta que la regla trataba de prohibir. Ser "preciso" significa especificar y, por ello, limitar las circunstancias en que la ley se aplica. El efecto de la fisura o laguna puede minimizarse volviendo a la norma deliberadamente demasiado incluyente: por ejemplo, estableciendo un límite de velocidad menor a la velocidad óptima en las condiciones normales de la conducción. Los costos de la inclusión

[9] Por ejemplo, si el costo esperado del castigo para quienes roban es de 10, pero quienes no roban afrontan un costo esperado del castigo de 3 debido a la vaguedad del estatuto del robo, el castigo esperado efectivo para el robo es de sólo 7. Este efecto de las normas vagas es similar al del error legal, discutido después en XXII.2, y al de la inclusión excesiva discutido antes, en VII.2.

[10] Véase, por ejemplo, Thornhill *vs*. State of Alabama, 310 U.S., 88, 97-98 (1940); Smith *vs*. California, 361 U.S., 147 (1959).

[11] Adviértase la analogía existente con el efecto fiscal de imponer una obligación de rescatar (véase antes VI.9): en ambos casos la amenaza de responsabilidad induce a la persona amenazada a mantenerse resguardada de las circunstancias en las que esa amenaza podría llevarse a cabo.

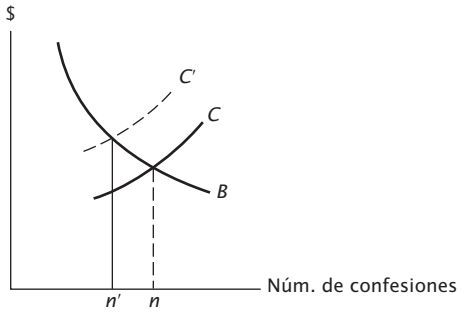
excesiva son menores entre más moderada sea la sanción y menores sean los costos de la elusión de la normas mediante una transacción. Por lo tanto, no debe sorprendernos observar: *1)* un uso mayor de la responsabilidad estricta en los contextos cuasidelictivos y contractuales que en los penales,[12] y *2)* muchas normas más o menos arbitrarias en los contextos legales de la propiedad y del comercio donde las normas pueden violarse mediante una transacción a bajo costo; se trata de contextos en los que puede darse un sentido económico al precepto de los abogados de que a menudo es más importante que se fije la ley aunque ello no se haga correctamente.

La claridad de una norma puede ser engañosa. Las normas generan presión para que se hagan excepciones, y la combinación de una norma con sus excepciones podría ser poco diferente, en términos prácticos, de un estándar, sobre todo si se permite que las excepciones *ad hoc* se multipliquen a fin de mejorar el ajuste entre una norma inicialmente rígida y un contexto social cambiante. Es más probable que un estándar produzca el equilibrio explícito de consideraciones opuestas (la contrapartida legal del análisis costo-beneficio); pero la administración de la norma podría generar el equilibrio indirectamente para resolver las disputas acerca del alcance y el significado de la norma y de sus excepciones. Porque, al decidir si una práctica particular debe clasificarse dentro de la norma o dentro de una de sus excepciones, los tribunales podrían apelar, abierta o encubiertamente, a un estándar que se considera subyacente o motivador de la norma.

Además, aunque una norma particular podría ser muy simple, es probable que todo el conjunto de las normas necesarias para cubrir el mismo campo que un estándar sea más complejo que el estándar. El estándar de la negligencia es más simple que el código de tráfico. Los estándares tienden a tener una fundamentación patente mientras que las normas tienden a ser arbitrarias, porque el elaborador de la norma está intercambiando compensaciones y pérdidas al sopesar la racionalidad sustantiva contra el costo de administración. Por ambas razones —complejidad y ausencia de justificación lógica a primera vista—, la regulación por la norma tiende a imponer costos de información mayores a las personas sujetas a la regulación que la regulación por el estándar. Pero la vaguedad de los estándares ejerce un efecto contrario y los costos de información de los reguladores pueden ser menores en un régimen de normas porque los hechos relevantes para la responsabilidad son menores y, así, los costos de la determinación legal son más bajos.

Vimos en el capítulo VII que cuando se acusa de una infracción penal de las leyes fiscales, el fiscal debe probar que el acusado conocía sus obligaciones en los términos de tales leyes, un tipo de prueba que no se requiere cuando se acusa de homicidio o de robo. El juicio implícito es que los costos de información del cumplimiento de normas complicadas o arbitrarias, pero en cualquier

[12] La negligencia es un estándar; la responsabilidad estricta es una norma.

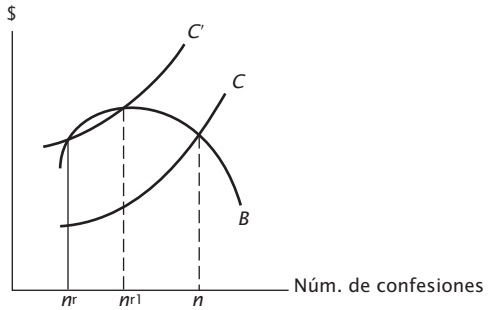


GRÁFICA XXI.1

caso no lógicamente justificadas a primera vista, son mayores que los del cumplimiento de estándares morales fundamentales. Un conjunto más complejo de normas podría reducir los costos reguladores y judiciales, pero disminuyendo al mismo tiempo el cumplimiento voluntario debido a los altos costos de información. Los recursos ahorrados por la dependencia administradora al determinar si la conducta conocida cumple o no con la ley podrían tener que gastarse en el incremento de las investigaciones del posible incumplimiento.

Algunos especialistas se muestran hostiles hacia el análisis económico del derecho porque creen que mina los derechos al disolver todas las cuestiones en un análisis costo-beneficio. Este problema acaba por relacionarse estrechamente con la elección entre normas y estándares. La crítica es que en el fondo el economista cree que toda norma jurídica tiene o debería tener excepciones que la transforman en un estándar, que en última instancia depende del equilibrio de los costos y de los beneficios. De modo que, por ejemplo, se acusa al economista de creer que la norma que prohíbe el uso de una confesión forzada como evidencia en un juicio penal contiene una excepción implícita: a menos que los beneficios de usarla superen los costos en un caso en particular.

El problema se ilustra en la gráfica XXI.1. El eje vertical es aquí el equivalente monetizado de los beneficios y los costos de las confesiones forzadas. El eje horizontal es el número de tales confesiones. *B* representa los beneficios marginales (en términos de una mayor certeza de que se condena al culpable) del uso de las confesiones forzadas como prueba en los juicios penales y se representa como descendente en la medida en que aumenta el número de las confesiones forzadas. *C* describe los costos marginales de tal uso (en las condenas erróneas, en el dolor y la humillación infligidos a la persona obligada a confesar, etc.) y se representa como creciente con el número de tales confesiones. La intersección, *n*, denota el número óptimo de las confesiones. Si sólo se permiten unas cuantas confesiones forzadas, la policía se limitará a los casos donde el exceso de los beneficios sobre los costos es muy grande, de modo que tenderá a evitar el uso de ese instrumento cuando los beneficios sean pequeños o los costos grandes.



GRÁFICA XXI.2

Supongamos que, debido a que desconfiamos de la capacidad de los tribunales para equilibrar costos y beneficios en los casos de confesiones y deseamos que se despejen todas las dudas en favor del acusado, establecemos una norma que prohíba toda coerción, excepto la más ligera (todo interrogatorio por parte de los custodios tiene algo de coercitivo), y eso sólo si no hay otro procedimiento para obtener una prueba esencial para condenar a quienes cometan delitos mayores. Esta norma, representada por la línea vertical continua en la gráfica XXI.1, reduce el número de las confesiones forzadas a n'. En efecto, hemos elevado la curva del costo a C' a fin de tomar en cuenta los costos ocultos o sutiles, con lo que sólo se permiten algunas confesiones forzadas, las que se encuentran en la región (reducida) a la izquierda de la intersección de la nueva curva de costos y la de beneficios que no se ha modificado.

Con todo, este ajuste deja intacta la premisa básica de la gráfica XXI.1: que un enfoque de reglas con excepciones, aunque distinto de un estándar, es por lo menos conmensurable con él.[13] Desestimar la premisa se representa en la gráfica XXI.2, en la que los costos y los beneficios de las confesiones forzadas se ilustran, plausiblemente, como una función de la cantidad de brutalidad ejercida para extraer la confesión. *B* tiene ahora una región creciente y una decreciente, porque cuando sólo se aplica una brutalidad pequeña se obtiene escasa información útil, mientras que en la medida en que se aplica una brutalidad mayor se obtienen durante cierto tiempo confesiones más valiosas, pero finalmente disminuye el valor de éstas cuando los acusados empiezan a confesar sólo para evitar un dolor insoportable. Sea que n^r denote el efecto de una norma que prohíbe el uso de cualesquiera confesiones, excepto las que se obtengan sin ninguna coerción o sólo con la coerción mínima implicada en el interrogatorio aplicado a una persona bajo custodia, por breve y nada amenazante que éste sea. Esta regla con excepciones no puede convertirse en un estándar, porque si aumentamos *C* hasta su intersección con *B*

[13] Véase, por ejemplo, Hilary Putnam, "Taking Rules Seriously", en Putnam, *Realism with a Human Face,* 193 (James Conant [comp.], 1990).

en n^r esto implica que las confesiones obtenidas con un mínimo de brutalidad —las confesiones que se encuentran a la izquierda de n^r, que en efecto la norma permite— están prohibidas porque sus costos (C') superan a sus beneficios, mientras que se permiten varias otras confesiones —las que se encuentran entre n^r y n^{r1}— que la norma prohíbe en efecto. En este ejemplo, el enfoque económico de las normas resulta en verdad incompatible con el enfoque no económico.[14] La decisión de cuál enfoque deberá tomarse es moral antes que económica.

Un punto final a destacar acerca de las reglas contra los estándares es que, situada entre aquéllas y éstos, se halla la "prueba multifactorial", que es la opción favorita de muchos jueces. Se trata generalmente de una lista abierta de factores que los jueces deben considerar al decidir, por ejemplo, si en un caso particular se ha de compensar a la parte triunfadora cubriendo sus cuotas legales, o si se debe fallar que se cometió infracción contra una marca registrada. Las pruebas tienden a ser de final abierto: pueden enlistar tres o siete o una decena de factores y, aun así, agregar que no se trata de una lista exhaustiva. El resultado de ello es que la prueba, pese al aura de precisión generada mediante la enumeración de factores específicos en busca de la consideración judicial, se convierte en algo mucho más cercano a un estándar que a una norma; y esto obedece adicionalmente a que los factores rara vez son ponderados, de manera que el juez se queda sumido en la incertidumbre sobre si, por ejemplo, en la prueba de siete factores, donde de un lado se alinean cuatro factores mientras del otro lado se alinean tres, debe decidir a favor de la parte que cuenta con más factores o a favor de la otra parte.

XXI.4. *STARE DECISIS*

La discusión de la sección anterior sugiere una razón importante, además del interés propio de los jueces (incluida la aversión al esfuerzo), para preferir la decisión de acuerdo con el precedente frente a la decisión tomada en cada juicio de nuevo sobre la base de primeros principios. El último enfoque sería el equivalente de las decisiones tomadas de acuerdo con un estándar amplio; tendría las desventajas discutidas en la sección anterior más los costos procedimentales asociados a la mayor incertidumbre que discutiremos en el capítulo siguiente. Por lo tanto, entre más amplio y particularizado sea el código estatutario de una nación, menor será el énfasis que podremos esperar en la *stare decisis*, como lo hicimos notar en el capítulo previo refiriéndonos a la práctica judicial en la Europa continental. Esperaríamos también que haya menos adhesión rígida al *stare decisis* entre más rápidamente esté cambiando

[14] Véase más sobre normas y estándares después en XXII.8; sobre las confesiones forzadas, véase XXX.3 después.

la sociedad. El cambio vuelve irrelevantes muchos precedentes, y erróneos a algunos de ellos. En ambas situaciones hay una depreciación rápida, pero en la primera el precedente deja simplemente de aplicarse y en la segunda debe ser revocado.

Un sistema de decisión de acuerdo con el precedente reduce los costos de la litigación al permitir que las partes de un juicio, y también el tribunal, utilicen información que ha sido generada (a menudo a un costo considerable) en juicios anteriores. Si en 20 juicios se ha sostenido que una señal eléctrica de cruzamiento es una precaución requerida en los cruces ferroviarios congestionados, la ganancia marginal del conocimiento de los costos y los valores relevantes que se obtiene por incurrir en el gasto de un juicio en el caso número 21 podría ser menor que ese gasto. Surge una norma del derecho común cuando sus premisas fácticas han sido validadas de tal modo por reiteradas verificaciones en la litigación que los gastos adicionales en pruebas y argumentaciones superarían al valor del conocimiento adicional producido. Pero, entre más rápidamente estén cambiando las condiciones sociales y económicas relevantes, mayor será el valor del conocimiento adicional y, por ende, menos valiosa será la adhesión al precedente.

No obstante, la *stare decisis* tiene un aspecto negativo: la "dependencia de la ruta" (véanse antes VIII.1 y X.11): el punto donde uno acabe podría depender del punto de donde uno parta, aun cuando, si no hubiera partido de donde lo hizo, un punto final diferente podría ser mejor.[15]

El ejemplo mejor conocido de la dependencia de la ruta en la literatura económica, aunque muy probablemente sea espurio, se refiere al teclado de la máquina de escribir. Según el historiador económico Paul David, el teclado se diseñó para limitar la rapidez de la escritura a fin de impedir el trabado constante de las teclas. El problema del trabado desapareció con el advenimiento de las máquinas de escribir eléctricas y el procesador de textos de la computadora, pero nos quedamos con el teclado antiguo porque serían prohibitivos los costos de que los fabricantes se pusieran de acuerdo sobre un nuevo teclado y de efectuar la "reconversión" de millones de personas que se habían capacitado con el teclado antiguo y se habían habituado a él.[16] Es de esperarse entonces que observemos la dependencia de la ruta cuando los cos-

[15] Adviértase la analogía que existe aquí con la cuestión de la adhesión rígida a la definición de los derechos de propiedad sobre la base del primero en tiempo-primero en derecho, discutida antes en III.5.

[16] Véase Paul A. David, "Clio and the Economics of QWERTY", 75 *Am. Econ. Rev. Papers & Proceedings,* 332 (mayo de 1985), y una crítica en S. J. Liebowitz y Stephen E. Margolis, "The Fable of the Keys", 33 *J. Law & Econ.,* 1 (1990); Liebowitz y Margolis, "Path Dependence, Lock-In, and History", 11 *J. Law, Econ. & Organization,* 205 (1995). Sobre una perspectiva general de la economía de la dependencia de la ruta, véase Stanley M. Besen y Joseph Farrell, "Choosing How to Compete: Strategies and Tactics in Standardization", *J. Econ. Perspectives*, primavera de 1994, 117. Véase también Lucian Arye Bebchuk y Mark J. Roe, "A Theory of Path Dependence in Corporate Ownership and Governance", 52 *Stan. L. Rev.,* 127 (1999).

tos de la transición sean elevados en relación con los beneficios del cambio, y tales costos tienden a ser elevados cuando la transición requiere un alto grado de coordinación.

La dependencia de la ruta es un fenómeno importante en el campo del derecho, y un ejemplo notable es la influencia (analizada en el capítulo XXXI del presente libro) del país donde se origina el sistema jurídico de una nación sobre las leyes vigentes en ella, pese a que su sistema pueda haber estado evolucionando de forma independiente durante centenares de años.[17] Hallamos más pruebas de la importancia que reviste la dependencia de la ruta en el derecho y la ley en que la convergencia de los sistemas jurídicos es mucho más lenta que la convergencia de las instituciones económicas. Por ejemplo, las leyes y las instituciones jurídicas de los diferentes estados de la Unión Americana difieren más que las prácticas y las instituciones económicas de los estados, y las diferencias son aún mayores y más misteriosas en una comparación entre países, aun cuando la comparación se limite a los países cuyos sistemas económicos y políticos, y los niveles de educación y de ingreso, son similares a los nuestros. Sin embargo, la razón de la dependencia de la ruta en el campo del derecho no es un problema de coordinación, como en el ejemplo del teclado de la máquina de escribir, sino el costo para las empresas y los individuos de la adaptación a un cambio repentino de la ley al cambiar las prácticas que habían adoptado y hasta la fecha seguían en apego a la ley antes de que ésta cambiara. Esto resulta especialmente claro en lo tocante a las leyes que establecen los derechos de propiedad o contractuales. Además, los jueces, así como los juristas en general, pueden estar tan privados de buenas fuentes de información para decidir los juicios novedosos o para reformar las instituciones legales a fin de que se ajusten a los cambios sociales que su método más eficiente para decidir los juicios y resolver las cuestiones del diseño institucional consistirá en seguir de cerca el precedente y ceder a las legislaturas el papel protagónico en lo que respecta a esforzarse porque la ley se adapte al cambio.

Lecturas recomendadas

Ehrlich, Isaac y Richard A. Posner, "An Economic Analysis of Legal Rulemaking", 3 *J. Legal Stud.,* 257 (1974).

[17] Adviértase asimismo la relación entre la dependencia de la ruta y las externalidades de la red, que se abordaron antes en X.11. Éstas hacen costosa la nueva entrada, pues el que apenas ingresa, al no tener una red, no puede ofrecer a los consumidores un producto igual de valioso a menos que incurra en el riesgo y el gasto de crear la red entera desde el comienzo, lo que probablemente precipitará una competencia ruinosa con el propietario titular de la red. Cualquier cosa que obstaculice la entrada reducirá las probabilidades del desarrollo de nueva tecnología si el valor de la nueva tecnología en relación con sus costos no excede las barreras a la entrada que el nuevo elemento entrante debe superar a fin de ser competitivo en el mercado a que desea acceder.

Fon, Vincy y Francesco Parisi, "On the Optimal Specificity of Legal Rules", 3 *J. Institutional Econ.,* 147 (2007).
Gennaioli, Nicola y Andre Shleifer, "The Evolution of Common Law", 115 *J. Political Econ.,* 43 (2007).
———, "Overruling and the Instability of Law", 35 *J. Comparative Econ.,* 309 (2007).
Jacobi, Tonja, "The Role of Politics and Economics in Explaining Variation in Litigation Rates in the U.S. States", 38 *J. Legal Stud.,* 205 (2009).
Kaplow, Louis, "A Model of the Optimal Complexity of Legal Rules", 11 *J. Law, Econ. & Organization,* 150 (1995).
Krehbiel, Keith, "Legislative Organization", *J. Econ. Perspectives,* invierno de 2004, 113.
Landes, William M. y Richard A. Posner, "Legal Precedent: A Theoretical and Empirical Analysis", 19 *J. Law & Econ.,* 249 (1976).
Miceli, Thomas J., "Legal Change: Selective Litigation, Judicial Bias, and Precedent", 38 *J. Legal Stud.,* 157 (2009).
O'Hara, Erin Ann, "Social Constraint or Implicit Collusion? Toward a Game Theoretic Analysis of *Stare Decisis"*, 24 *Seton Hall L. Rev.,* 736 (1993).
Priest, George L., "Measuring Legal Change", 3 *J. Law, Econ. & Organization,* 193 (1987).
Production of Legal Rules, parte 2 (Francesco Parisi [comp.], 2011).
Rasmusen, Eric, "Judicial Legitimacy as a Repeated Game", 10 *J. Law, Econ. & Organization,* 63 (1994).
Stearns, Maxwell L., *Constitutional Process: A Social Choice Analysis of Supreme Court Decision Making*, cap. 4 (2000).

Problemas

1. ¿Por qué es incorrecto esperar que el volumen de la litigación en un campo del derecho crezca como una función simple de las actividades (número de divorcios, delitos, accidentes automovilísticos, etc.) que originan la litigación?

2. ¿Por qué sería de esperarse que el Parlamento de Inglaterra desempeñara un papel mayor en la producción de normas de derecho, en relación con los tribunales ingleses, que el Congreso de los Estados Unidos en relación con los tribunales federales?

3. Discuta los pros y contras de la revocación prospectiva de decisiones anteriores a la luz del análisis de este capítulo.

4. ¿Sugiere este capítulo una posible teoría económica del principio constitucional de la protección igual de las leyes?

5. Discuta los costos y beneficios de la complejidad en el derecho fiscal. Véase Louis Kaplow, "Accuracy, Complexity, and the Income Tax", 14 *J. Law, Econ. & Organization,* 61 (1998).

XXII. PROCEDIMIENTO CIVIL Y PENAL

XXII.1. Las metas económicas del procedimiento

El objetivo de un sistema procedimental, desde el punto de vista económico, consiste en minimizar la suma de dos tipos de costos. El primero es el costo de las decisiones judiciales erróneas. Supongamos que el costo esperado de un tipo de accidente particular es de 100 dólares y el costo de evitarlo para el victimario potencial es de 90 (supondremos que el costo de evitarlo para la víctima es mayor de 100). Si el victimario potencial está sujeto a un estándar de negligencia o de responsabilidad estricta, evitará el accidente… si suponemos que el estándar se administra correctamente. Pero supongamos que en 15% de los casos en los que ocurre un accidente puede esperarse que el victimario evite la responsabilidad debido a las determinaciones fácticas erróneas del sistema procedimental. Entonces bajará a 85 dólares el costo esperado del accidente para el victimario, y, dado que esto es menor al costo de evitarlo para él (90), no se prevendrá el accidente. El resultado será una pérdida social neta de 10 dólares. ¿O no?[1]

No debemos omitir el costo de operación del sistema procedimental. Supongamos que para bajar la tasa del error de la no imposición de responsabilidad, de 15% a menos de 10%, se requeriría una inversión adicional en el sistema procesal de 20 dólares por accidente.[2] Entonces tendríamos que tolerar la probabilidad de error de 15%, porque el costo del error (10 dólares) es menor al costo necesario para eliminarlo (20). No importa si el propósito de la regla sustantiva aplicada al caso en cuestión es promover la eficiencia económica. Siempre y cuando un valor del dólar pueda ubicarse en los beneficios de una decisión correcta, los costos de tomar semejante decisión se pueden comparar con los costos del error.

Este tipo de comparación está implícito en la opinión de la Corte Suprema sobre Mathews *vs.* Eldridge,[3] donde se sostuvo que, al decidir cuánto proceso se debe a alguien que se queja de que el gobierno lo ha privado de su propie-

[1] Véase Charles Alan Wright y Mary Kay Kane, *Law of Federal Courts* (6ª ed., 2002); Stephen C. Yeazell, *Civil Procedure* (6ª ed., 2004).

[2] El sistema procedimental sólo se emplea cuando un accidente origina un juicio, pero dividiendo el costo del sistema por el número de los accidentes obtenemos una cifra del costo que puede compararse con el costo medio de un accidente.

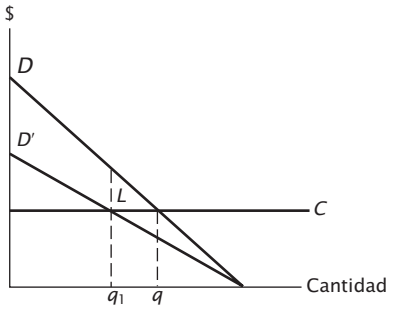
[3] 424 U.S., 319 (1976). Véase también Sutton *vs.* City of Milwaukee, 672 F.2d, 644, 645 (7th Cir., 1982); y una crítica del enfoque de la Corte Suprema en Jerry L. Mashaw, "The Supreme Court's Due Process Calculus for Administrative Adjudication in Mathews *vs.* Eldridge. Three Factors in Search of a Theory of Value", 44 *U. Chi. L. Rev.,* 28, 47-49 (1976).

dad, los tribunales debieran considerar el valor de la propiedad, la probabilidad de la privación errónea porque se omitió la salvaguardia procedimental particular que se buscaba y, en fin, el costo de la salvaguardia. En términos de la Fórmula de Hand, se niega el debido proceso cuando $B < PL$, donde B es el costo de la salvaguardia procedimental, P es la probabilidad del error si la salvaguardia se niega y L es la magnitud de la pérdida si el error se materializa.

Por supuesto, al igual que en el caso de la Fórmula de Hand misma, pocas veces es posible cuantificar estos términos (o por lo menos no se hacen esfuerzos en tal sentido).[4] Pero la fórmula es valiosa incluso cuando se usa de manera cualitativa antes que cuantitativa. Supongamos, por ejemplo, que la cuestión consista en saber si el propietario de un automóvil aparentemente abandonado debiera ser notificado y recibir una oportunidad de una audiencia, *antes* de que el automóvil sea remolcado y vendido como chatarra. No es trivial la probabilidad de que el automóvil no estuviera realmente abandonado, sino que se hubiera descompuesto o se lo hubieran robado, y el costo de una audiencia es moderado en relación con el valor del automóvil; por lo tanto, como han sostenido la mayoría de los tribunales, el propietario debiera tener derecho a la audiencia. Pero supongamos que no estamos hablando de automóviles abandonados sino estacionados en lugares prohibidos. Dado que los automóviles no están a punto de ser destruidos, la privación (L) es mucho menor que en el caso del automóvil abandonado. La probabilidad del error es también mucho menor, porque de ordinario es simple determinar que un automóvil está estacionado en un lugar prohibido. Y es muy elevado el costo de una audiencia antes de la privación: si el propietario debe ser notificado antes de que el automóvil sea remolcado, lo moverá y se eliminará el efecto disuasivo del remolque. En consecuencia, los tribunales sostienen que el debido proceso no requiere una audiencia antes de la privación en el caso de los automóviles estacionados en lugares prohibidos.

¿Qué ponderación se le debe dar al hecho de que el pago de una multa por estacionar el carro en el lugar prohibido es un pago de transferencia, en tanto que un carro abandonado será con toda probabilidad arrumbado en los basureros? No tanta como uno pudiera pensarlo, debido a que si la tasa de error en la emisión de multas para autos estacionados en el lugar indebido fuese sustancial habría efectos adversos sobre la conducción de automóviles y sobre la práctica de estacionarlos. Por ejemplo, si la emisión de multas fuese aleatoria habría más ocupación ilegal de lugares de estacionamiento, puesto que la probabilidad gradual de recibir una multa si uno dejara el auto en un lugar prohibido sería nula.

[4] Véase un ataque a la cuantificación en Van Harken *vs.* City of Chicago, 103 F.3d, 1346 (7th Cir., 1997).



GRÁFICA XXII.1

XXII.2. LOS COSTOS DEL ERROR EN LOS CASOS CIVILES

La gráfica XXII.1 representa los efectos de las determinaciones judiciales erróneas. *D* representa el valor de una unidad de equipo de seguridad, en términos de los costos de los accidentes evitados, como una función de la cantidad del equipo comprado; en otras palabras, *D* es el producto marginal del equipo de seguridad. C representa el costo del equipo. Se maximiza el valor cuando se lleva la compra de equipo de seguridad hasta el punto q, en que se intersecan *D* y *C*. Si la industria es plenamente responsable del costo de los accidentes, *D* se convierte en la demanda de equipo de seguridad por parte de la industria y se instalará la cantidad óptima. Pero si debido al error en el sistema jurídico puede esperarse que la industria sea responsable de sólo $p\%$ de los costos de los accidentes que inflige, su demanda de equipo de seguridad bajará a D' (igual a $p \times D$), y sólo comprará la cantidad q_1 de equipo de seguridad, lo que generará una pérdida social de *L*. El efecto del error sobre el comportamiento de la industria es el mismo que el de un impuesto a los ingresos brutos (véase antes XVIII.3).

Pero el análisis es incompleto en dos sentidos importantes. Primero, cualquier factor que cause errores en favor de la industria en algunos casos (testigos que mienten, por ejemplo) causará errores en favor de las víctimas de los accidentes en otros casos, y esto desplazará a D' hacia la derecha. Sin embargo, es probable que, en total, D' se encuentre debajo de *D*. Todos los errores en favor de la industria mueven a D' hacia abajo; sólo algunos errores contra la industria la mueven hacia arriba. Un error que incremente la responsabilidad de la industria en los casos en que ha sido negligente o culpable por alguna otra razón (la víctima del accidente convence al tribunal de que sus daños son mayores de lo que en realidad son) tenderá a volver al equipo de seguridad más valioso para la industria y a elevar en consecuencia a D'. Pero si el error se traduce en una responsabilidad cuando el equipo de seguridad adicional no habría reducido la responsabilidad de la industria (es

posible que el accidente no haya ocurrido en realidad, o que haya sido causado por alguien que no sea miembro de la industria), ésta no tendrá ningún incentivo para comprar equipo de seguridad adicional, y D' no aumentará.

Segundo, es probable que el costo social de una disminución del incentivo de la industria para evitar los accidentes se vea parcialmente contrarrestado por el incremento del incentivo de las víctimas para evitar los accidentes. El efecto de un desplazamiento de D' hacia abajo es un aumento del costo esperado y no compensado del accidente para las víctimas, y, por lo tanto, un aumento de su incentivo para prevenir los accidentes. En estos casos, aunque evitar el accidente por parte de las víctimas será menos eficiente que si se evita por parte de los victimarios (pues de otro modo la responsabilidad o la responsabilidad parcial habrían sido adjudicadas a la víctima mediante la aplicación de la doctrina de la negligencia contributiva o de la negligencia comparada) y no hubiera error alguno, existiría alguna compensación.

El error legal puede alterar las reglas *de facto* sustantivas. Recordemos que la doctrina contractual de la imposibilidad trata de desplazar por completo ciertas pérdidas del promitente al promisario (véase antes IV.5). La posibilidad del error judicial genera cierta probabilidad de que el promitente se quede con la pérdida, de modo que considerada *ex ante* es la doctrina de compartición de la pérdida antes que de traslado de la pérdida. El error elimina el carácter dicotómico de los fallos judiciales.

XXII.3. La adjudicación anticipatoria. Origen de las órdenes preliminares, los fallos declaratorios y las opiniones consultivas

La importancia de ponderar los errores por sus costos se pone de manifiesto también en los casos en que se pide a los tribunales que fallen sobre la base de una información menos completa que la que se tendría si se demorara el fallo. Los trueques esenciales de ventajas y pérdidas se observan sintéticamente en la fórmula siguiente para decidir si deberá otorgarse o negarse una orden preliminar: otorgarla si, y sólo si, $P(Hp) > (1 - P)\, Hd$, donde P es la probabilidad de que el demandante gane el juicio completo por sus méritos (y, por lo tanto, $1 - P$ es la probabilidad de que gane el demandado), Hp es el daño irreparable que sufrirá el demandante si no se otorga una orden preliminar para mantener el *statu quo* mientras se tramita el juicio, y Hd es el daño irreparable que sufrirá el demandado si se otorga la orden preliminar. (Si el daño no es irreparable —si puede ser reparado por un remedio que se ordene al final del juicio—, no hay ninguna razón para apresurar las cosas y, al proceder de ese modo, incurrir posiblemente en costos de errores adicionales. Equivaldría a apresurar la construcción de una casa, a un costo adicional, a pesar de que el comprador no tuviese ningún deseo de ocuparla pronto.)

El riesgo de error es grande porque se está pidiendo al juez que falle deprisa, sobre la base de una información incompleta. Con todo, el juez puede minimizar los costos del error esperado comparando los costos del error ponderados (los dos lados de la desigualdad) de las partes. Supongamos que el demandante tiene una probabilidad de 60% de ganar tras un juicio completo, que el riesgo del error de negar su petición de una orden preliminar es también de 60% y que el riesgo del error si el juez otorga la orden debe ser por lo tanto de 40%. Pero supongamos además que si se niega la orden el demandante sufrirá un daño irreparable de 50 dólares, mientras que si se otorga la orden el demandado sufrirá un daño irreparable de 100. Entonces el costo esperado del error es mayor para el demandado que para el demandante (40 contra 30 dólares) y la orden debiera negarse. A veces los tribunales utilizan un enfoque que es una aproximación a la fórmula presentada antes.[5]

Supongamos que la parte que pide un auxilio judicial inmediato no es el demandante sino un demandado potencial que busca un fallo declaratorio, es decir, un fallo en el sentido de que su conducta potencial no viola los derechos legales del demandante potencial (esto es, el demandado en la acción del juicio declaratorio). La ley distingue claramente (como debe ser desde un punto de vista económico) entre los casos en que las partes no tienen una disputa efectiva y los casos en que sí la tienen. Si no hay disputa efectiva, y el demandado potencial sólo está buscando un consejo acerca de sus obligaciones legales, los tribunales no concederán el auxilio declaratorio. Estarían abrumados con los casos planteados por personas que buscan un consejo más barato y más autorizado que el que podrían obtener de abogados privados; y carecerían de la información que resulta de una disputa efectiva.

Pero supongamos que las partes tienen una disputa efectiva, como ocurre cuando se pide a una compañía aseguradora de responsabilidades que defienda un juicio promovido contra su asegurado y la cobertura es dudosa. Si la compañía niega la defensa y más tarde se descubre que ha violado el contrato de seguro, podría tener que pagar daños punitivos al asegurado (véase antes VIII.2). Por lo tanto, la compañía de seguros podría tratar de obtener, antes del juicio contra su asegurado, un fallo declaratorio en el sentido de que no hay cobertura y, por lo tanto, no está obligada a defender al asegurado. El requisito de que haya en verdad una disputa limita el número potencial de los casos en los que se busca el auxilio declaratorio. En efecto, en el ejemplo el fallo declaratorio probablemente habrá de eliminar un juicio futuro del asegurado contra la compañía aseguradora, y por ello la disponibilidad del auxilio

[5] Roland Machinery Co. *vs.* Dresser Industries, Inc., 749 F.2d, 380, 387-388 (7th Cir., 1984); John Leubsdorf, "The Standard for Preliminary Injunctions", 91 *Harv. L. Rev.*, 525 (1978). La fórmula puede reordenarse en una forma que la vuelva más comprensible de entrada para algunos lectores: $P/(1 - P) > H_d/H_p$. En palabras, debiera otorgarse la orden preliminar si, pero sólo si, la razón de las probabilidades de triunfo del demandante a las del demandado supera a la razón del daño irreparable del demandado con respecto al del demandante.

declaratorio no habrá incrementado el número de casos siquiera por uno. En virtud de que la cobertura depende por entero de una comparación del lenguaje del contrato de seguros con el lenguaje del juicio contra el asegurado, el tribunal tiene tanta información acerca de los méritos de la disputa entre la compañía de seguros y su asegurado como la que tendría si se pospusiera la resolución de la disputa hasta que se promoviera un juicio del asegurado contra la compañía de seguros. Adviértase también que en virtud de que la obligación de defender a expensas de la compañía aseguradora es independiente del fallo del juicio promovido contra el asegurado, el hecho de que el asegurado pueda ganar en el juicio de responsabilidad en su contra no resolverá su disputa con la compañía aseguradora. Por lo tanto, no hay nada que ganar con la demora.

Cuando no hay todavía una disputa efectiva entre las partes, decimos que el solicitante está buscando una "opinión consultiva". Las reglas de legitimación (véase antes XX.2) prohíben que los tribunales federales emitan opiniones consultivas. Sin embargo, algunos estados de la Unión Americana permiten que sus cortes supremas emitan opiniones consultivas en lo tocante a la constitucionalidad de la legislación estatal que se ha promulgado, pero no ha entrado en vigor todavía. Esta práctica ("revisión abstracta": a menudo el único tipo de revisiones que realizan los tribunales constitucionales extranjeros) evita los costos asociados al hecho de poner en vigor una ley que habrá de declararse inconstitucional más tarde. Sin embargo, este beneficio se compra a cierto precio en términos de la pérdida de información acerca de la administración y los efectos reales de la ley, una pérdida que podría conducir a invalidaciones erróneas de algunas disposiciones constitucionales. La pérdida potencial es menor al nivel estatal que al nivel federal porque las constituciones estatales pueden enmendarse para corregir una interpretación judicial errónea con mucha mayor facilidad que en el caso de la Constitución federal. Ésta podría ser la razón de que encontremos opiniones consultivas a nivel estatal pero no a nivel federal.

XXII.4. El juicio sumario y las peticiones de desestimación

Una orden judicial no pone término a la litigación tomada como una cuestión formal, pero a menudo tal es su efecto práctico. Otros tipos de procedimientos previos al juicio son de carácter final, como el de resolver una demanda sobre el juicio sumario o el de conceder una moción para desestimar una acción debido a la incapacidad de expresar un reclamo. Una moción para desestimar se antepone al momento en que las partes se involucran en el descubrimiento previo al juicio y, así, en caso de ser concedida, el caso se resuelve a un bajo costo para las partes o para el tribunal. Una moción para el juicio sumario se levanta usualmente una vez que se hubo finalizado el descubrimiento previo

al juicio, o si se levanta antes es probable que el tribunal difiera la emisión del fallo hasta que dicho descubrimiento se haya llevado a cabo. En el caso de que la evidencia obtenida en el juicio sumario no logre, si es presentada en el juicio, orillar a la formulación de un veredicto a favor de la parte en movimiento, la moción es negada. Por lo habitual, la única diferencia entre esa evidencia y la evidencia presentada ante el tribunal es que la primera se limita a los documentos mientras que la segunda es oral.

El problema con el procedimiento del juicio sumario consiste en la frecuente asimetría de las cargas que constituyen el descubrimiento. El demandante por lo general quiere contar con más descubrimientos que el demandado, porque aquél quiere buscar los registros de éste en busca de pruebas de las malas acciones cometidas por éste. En una era de descubrimientos electrónicos, cuando a menudo las corporaciones tienen en su haber millones de e-mails y otras comunicaciones electrónicas en sus archivos, al lado de extensos registros en papel, el costo para el acusado de satisfacer las demandas de descubrimiento del demandante pueden ser en verdad considerables, y, si no hay para el demandante costos similares por cubrir las demandas de descubrimiento hechas por el acusado, la asimetría surgida en los costos de descubrimiento puede inducir al demandado a convenir en arreglar el caso en términos favorables al demandante aun cuando el caso tenga un escaso mérito.

Respondiendo a esta preocupación por el descubrimiento abusivo, en fecha reciente la Corte Suprema relajó el estándar para la concesión en casos federales de una moción para desestimar una acción debido a la incapacidad de expresar un reclamo.[6] La prueba para validar esa desestimación solía ser si había un estado de los hechos cualquiera, consistente con las alegaciones de la queja, el cual, de ser probado, establecería el derecho del demandante a un juicio; si la respuesta era positiva, el caso no podía ser descartado. La nueva prueba consiste en ver si la queja realiza una demostración "plausible" de mérito; si no lo hace, el caso debe descartarse. Para establecer la plausibilidad se requieren alegatos más cabales, y, por consiguiente, una investigación previa a la queja más extensa que corre a cargo del demandante, y así genera una mayor simetría entre los costos de litigación del demandante y el acusado. Si el demandante muestra que no puede conducir una investigación siquiera mínimamente adecuada sin un descubrimiento limitado, el juez puede dar cabida a ese descubrimiento, y, mientras tanto, diferirá el fallo sobre la moción del acusado de que se proceda a desestimar.

[6] Véanse Ashcroft *vs.* Iqbal, 556 U.S., 662 (2009); Bell Atlantic Corp. *vs.* Twombly, 550 U.S., 544 (2007).

XXII.5. La decisión de transigir o ir a juicio

Podría parecer que el hecho de que las controversias se litiguen en lugar de transigirse viola el principio de que, cuando los costos de transacción son bajos, las partes transigirán voluntariamente si es posible de una manera mutuamente benéfica. En efecto, pocos casos se van a juicio; un estudio descubrió que sólo 2% de las reclamaciones por accidentes automovilísticos van a juicio,[7] y otro descubrió que sólo se instruyen 4% del total de los casos civiles de los tribunales estatales.[8] Por supuesto, muchos casos se resuelven antes del juicio, o incluso son abandonados. Comoquiera que sea, la mayoría de ellos llegan a un arreglo (incluidos casi todos los casos penales), como lo pronosticaría la teoría económica, pero todavía tenemos que explicar la pequeña fracción de los que se van a juicio.

Como ocurre con cualquier contrato, una condición necesaria para el éxito de las negociaciones es que haya un precio al que ambas partes concluyen que el acuerdo aumentaría su bienestar. Por lo tanto, las negociaciones para el arreglo fracasarán, y seguirá una litigación, si el precio mínimo que el demandante esté dispuesto a aceptar como solución negociada de su reclamación es mayor al precio máximo que el demandado esté dispuesto a pagar como satisfacción de esa reclamación. Por ejemplo, si el demandante no se arregla por menos de 10 000 dólares y el demandado no lo hace por más de 9 000, fracasarán las negociaciones para el arreglo.

La existencia de un área de traslape entre los términos mínimos de las partes, o precios de reserva —en otras palabras, un intervalo de arreglo—, es una condición necesaria para un arreglo, pero no es una condición suficiente. Las negociaciones para el arreglo son un ejemplo clásico del monopolio bilateral. El demandante puede transigir sólo con el demandado, y éste sólo puede transigir con el demandante, y ambos desean quedarse con la mayor parte posible del excedente generado por el arreglo sobre la litigación. En efecto, entre mayor sea el intervalo del acuerdo, las partes podrán ganar más con una negociación dura, y más probable será (aparentemente) que las partes acaben litigando porque no pueden ponerse de acuerdo sobre la división del excedente disponible. Pero hay algunos factores compensatorios: por definición, un intervalo más grande contiene más puntos de beneficio mutuo. Además, entre mayor sea el intervalo, menos costará a las partes determinar que un arreglo

[7] H. Laurence Ross, *Settled Out of Court: The Social Process of Insurance Claims Adjustments,* 179, 216 (2ª ed., 1980). Véase también Patricia Munch Danzon y Lee A. Lillard, "Settlement Out of Court: The Disposition of Medical Malpractice Claims", 12 *J. Legal Stud.*, 345, 365 (1983) (se litigó en menos de 10 por ciento).

[8] Brian J. Ostrom y Neil B. Kauder, *Examining the Work of State Courts,* 1994, en 28 (1996). La cifra correspondiente para los tribunales federales es de sólo 1.4%. Judicial Business of the United States Courts, 182 (Administrative Office of the U.S. Courts, 2005).

—algún arreglo— es lo más conveniente para ambas, y mayor será la pérdida potencial de no llegar a un acuerdo. No obstante, si bien, haciendo un balance, la dimensión del intervalo del arreglo y la probabilidad de llegar a un acuerdo bien pueden estar positivamente correlacionados, entre más grande sea el intervalo, será probablemente más larga la negociación del arreglo y, por tanto, éste será más caro en términos del tiempo de los abogados o de otros negociadores.

Este análisis deja un poco en el misterio al *proceso* del arreglo ("llegar al sí"). Los abogados creen tradicionalmente que si se sugiere al contrario un arreglo se señalará cierto pesimismo acerca de nuestras perspectivas de triunfo si el caso se va a juicio. Esto bien puede ser incorrecto. La primera oferta de arreglo a menudo puede tener lo que los economistas conductuales denominan un "efecto de anclaje": la parte opositora sentirá el "tirón" de la oferta al decidirse sobre una contraoferta, y de esta manera es probable que el arreglo último se aproxime más a la oferta original que si el oferente no la hubiera hecho en primer lugar.

Empero, si ninguna de las partes asume la iniciativa en las negociaciones de un arreglo, ¿cómo podrá empezar a haber negociaciones? El juez puede tomar la delantera, impeliendo a las partes a discutir el arreglo. Las reglas de los tribunales pueden incluso requerir negociaciones en pos de un acuerdo. Por supuesto, las negociaciones pueden parar en un fracaso. El arreglo podría parecer especialmente improbable si las partes, en virtud de que tienen una información diferente acerca de la solidez de sus posiciones respectivas, no convienen en el resultado probable de la litigación. Aun así, la divergencia de sus pronósticos acerca del resultado no es necesariamente fatal; todo depende del costo de la litigación en relación con el costo del arreglo. Además, el propio proceso de arreglo tenderá casi seguramente a reducir la divergencia de las estimaciones. Cada una de las partes tratará de volver más creíble su oferta o su demanda apoyándola con información favorable para su posición, y el hecho de que no se provea tal información podría ser igualmente revelador acerca de la fuerza o la debilidad de la posición de una de las partes (pronto volveremos a ocuparnos de este punto).

Pero, aun si no hubiera ninguna divergencia entre las estimaciones de las partes acerca del resultado probable en el juicio, la existencia de un *intervalo* significa que no hay ningún precio único para el arreglo y, en tal situación, las partes escogerán a menudo un punto medio del intervalo ("partirán la diferencia"). Esto economiza los costos de transacción y es una solución especialmente atractiva para un juego de negociación si ninguna de las partes tiene alguna razón para pensar que él es el negociador más fuerte.

En el sistema estadunidense (seguido por algunos otros países), donde el perdedor (aun cuando con algunas excepciones) no paga los costos del litigio de la parte ganadora, la ganancia neta esperada por el demandante es el monto del fallo que lo declare ganador descontado por su estimación de la

probabilidad de que gane, menos sus costos de litigación. La pérdida esperada por el demandado es el monto del fallo en caso de que pierda descontado por su estimación de la probabilidad de perder (o, dicho de otro modo, su estimación de la probabilidad de que gane el demandante), más sus costos de litigación. Si la ganancia esperada del litigio por el demandante es de 10 000 dólares, no se arreglará por menos de esa cantidad a menos que sienta aversión por el riesgo, una complicación que pospondremos por ahora; y si el demandado espera perder sólo 9 000 si la disputa se va a juicio, no se arreglará por más de esa cantidad (con la misma calificación). Además, la mejor oferta de arreglo se ajustará hacia arriba por el demandante y hacia abajo por el demandado, para reflejar los costos del arreglo. Si tales costos son, digamos, de 500 dólares para cada una de las partes, la demanda mínima del demandante será de 10 500 en el ejemplo anterior, y la oferta máxima del demandado será de 8 500 dólares.

En la desigualdad (1) se resume la condición necesaria para que ocurra la litigación. J es el monto del fallo si gana el demandante. P_p es la probabilidad de que gane el demandante estimada por él mismo, y P_d es la estimación de esa probabilidad por parte del demandado. C y S son los costos de la litigación y del arreglo para cada parte, respectivamente. Éste es un modelo muy simple porque supone que ambas partes son neutrales al riesgo y que los intereses en el caso, los costos de la litigación y los costos del arreglo son los mismos para ambas partes;[9] algunos de estos supuestos se relajarán más adelante.

La condición necesaria para la litigación,

$$P_pJ - C + S > P_dJ + C - S, \tag{1}$$

puede rescribirse como

$$(P_p - P_d)J > 2(C - S). \tag{2}$$

Si las partes convienen en la probabilidad de que el demandante ganará en caso de una litigación, el miembro izquierdo de (2) será 0 y el caso se arreglará, porque la litigación es más costosa que el arreglo; *a fortiori* se arreglará si una de las partes es más pesimista que la otra, de modo que $P_p - P_d$ es negativa. En general, entonces, sólo ocurrirá la litigación si ambas partes son optimistas acerca del resultado de la litigación.

Debemos considerar *por qué* divergen los litigantes en sus estimaciones del resultado probable si la litigación se va a juicio. Una razón es que cada parte tiene información privada, es decir, información no disponible para la otra parte, concerniente a ciertas facetas del caso, como el cuidado de sí o su

[9] También supone un resultado dicotómico de la litigación (alguna J fija o nada), y que los costos de la litigación y del arreglo son exógenos (es decir, no se ven afectados por otros términos de la fórmula). Este último supuesto se relaja después en XXII.11.

propio perjuicio. Ninguna de las partes puede comunicar fácilmente esa información a la otra en una forma creíble, convirtiendo así la información privada en pública, dado el valor táctico del secreto; su revelación descubriría debilidades que la otra parte podría explotar en las negociaciones del arreglo o en el juicio.

La desigualdad (2) pone de relieve la observación importante de que, en igualdad de las demás circunstancias, entre mayores sean los intereses en juego, más probable será que el caso sea litigado (es decir, más probable será que la desigualdad se deba satisfacer). La explicación justificada lógicamente a primera vista es que cuando los intereses son pequeños las ganancias potenciales de la litigación, tal como las perciben las partes, son también pequeñas y tienden a estar dominadas por costos de la litigación mayores que los de un acuerdo. Una consideración parcialmente compensatoria es que los casos mayores atraen a mejores abogados, quienes serán más hábiles para pronosticar el resultado de la litigación, de modo que disminuye $P_p - P_d$.

Cambiemos algunos de los supuestos del modelo. Supongamos que:

1. Los intereses del caso no son los mismos para ambas partes; es posible que las partes tengan tasas de descuento diferentes del valor futuro al valor presente, lo que hará que sus *J* diverjan. O quizás una de las partes anticipe beneficios futuros a partir del valor del precedente de ganar. Entonces *J* será para esta parte, en efecto, sólo el inicio de una corriente de beneficios. La cuestión decisiva es el monto de la divergencia de las *J*. La desigualdad (1) implica que, si la *J* del demandante es menor, la litigación es menos probable que si las dos *J* son del mismo tamaño, mientras que la litigación es más probable si la *J* del demandado es menor. Un arreglo puede ser imposible si una de las partes o ambas, poniendo el ojo en disputas futuras o con una tercera parte, desean desarrollar o mantener una reputación de negociante duro. Una o ambas pueden incluso considerar el costo del arreglo más alto que el costo de la litigación cuando el primero incluye una mayor probabilidad de disputas costosas en el futuro.
2. Las partes no son neutrales al riesgo. Si ambas partes sienten aversión por el riesgo, se reducirá la probabilidad de la litigación. El arreglo es un determinado desenlace, la litigación es su equivalente incierto. Los adversos al riesgo prefieren la certidumbre, así que aceptarán unos términos del arreglo algo menos favorables (sobre una base estricta de valor esperado) que el litigio. Esto aumenta el intervalo de la negociación y por ello la probabilidad de un arreglo.
3. Los costos de litigación y de arreglo de las partes no son realmente fijos, sino que varían con los intereses o, más realistamente aún, contienen un componente fijo y otro variable. Hay un gasto mínimo en la litigación o el arreglo de un caso, representado por *C* y *S* en las desigualdades

(1) y (2); y es probable que sea aproximadamente el mismo para ambas partes en la mayoría de los casos. Pero, fuera de eso, las partes gastarán más en la litigación entre más tengan que ganar: la litigación es una inversión además de un gasto. Así que es probable que, entre mayor sea *J*, más gaste cada parte en la litigación, porque todo aumento de P_p (para el demandante) o de P_d (para el demandado), provocado por un gasto adicional en la litigación, conferirá una ganancia esperada mayor entre mayor sea *J*. Presumiblemente, sin embargo, este componente variable del gasto en la litigación aumenta con menor rapidez que *J;* por ejemplo, podría aumentar como la raíz cuadrada de *J*.

Estos refinamientos, especialmente *2)* y *3)*, complican nuestro pronóstico anterior de que un aumento de los intereses en juego reducirá la tasa de los arreglos. Los intereses mayores aumentan el riesgo de la litigación al expandir la varianza de los resultados posibles, y, entre más riesgosa sea la litigación, más desearán el arreglo las partes que sienten aversión por el riesgo. Más importante aún, un aumento de los intereses evoca un aumento de los costos esperados de la litigación, y supongamos plausiblemente que ese aumento es mucho mayor que el aumento de los costos esperados del arreglo. No es mucho más costoso el arreglo de un caso grande que el de uno pequeño, pero resulta mucho más costosa su litigación. Por lo tanto, los intereses mayores vuelven al arreglo un sustituto de la litigación más barato aún.

Otra complicación presiona en la misma dirección. Cuando se arregla un caso por dinero, el acuerdo del arreglo no se presenta de ordinario ante el tribunal y las partes pueden mantener en secreto los términos del arreglo si así lo desean. Esto se convierte en un incentivo para que los demandados arreglen los casos grandes a fin de evitar un juicio público que alentaría otros juicios promovidos por personas situadas en una posición similar a la del demandante victorioso. El demandante aceptará en general con agrado tal disposición de confidencialidad porque ello permitirá un arreglo más rápido y mayor.

¿Debieran los tribunales hacer cumplir tales disposiciones? Supongamos que se presenta una secuencia de juicios similares contra el mismo demandado. Los demandantes que actúan y arreglan antes se beneficiarán del acuerdo de mantener en secreto los términos del arreglo, por la razón anotada antes; pero los demandantes que actúan más tarde resultan perjudicados porque si hubieran conocido los términos de los arreglos anteriores habrían podido hacer una estimación más correcta del valor de sus propias reclamaciones.[10] Así pues, haciendo un balance, no resulta claro si los demandantes ganan o pierden por dar lugar a disposiciones de confidencialidad; los arreglos confidenciales, a la vez que impiden el acceso a la información que otros deman-

[10] Véase Andrew F. Daughety y Jennifer F. Reinganum, "Informational Externalities in Settlement Bargaining: Confidentiality and Correlated Culpability", 33 *RAND J. Econ*., 587 (2002).

dantes podrían emplear para presionar en pos de buenos arreglos, incrementan la dimensión de los arreglos que los demandantes convienen en mantener en secreto.

Hasta aquí hemos venido suponiendo que hay un demandante y un demandado. Consideremos el ejemplo de un demandado que se enfrente a una serie de demandantes. Supongamos que éstos difieren en cuanto al grado de su optimismo acerca de sus probabilidades de ganar si se van a juicio. O podrían tener estimaciones diferentes acerca del monto del fallo. En cualquier caso, el valor esperado de la litigación y, por ende, el precio mínimo del arreglo variarán entre los demandantes. En tal situación, el demandado podría reducir el costo agregado de la litigación para él ofreciendo un acuerdo de arreglo que incluya una cláusula de la nación más favorecida (véanse X.1 antes y XXXI.2 después), por la que el demandado promete a cada uno de los demandantes que, si más adelante logra una mayor cantidad con otros demandantes, aumentará su pago del arreglo al mismo nivel.[11] Supongamos que el demandado se arregla sobre esta base con cuatro de cinco demandantes por 2 millones de dólares cada uno. Entonces, si se arreglara con el último demandante por 2.5 millones, el arreglo no le costaría 2.5 millones sino 4.5 millones de dólares, porque tendría que pagar una suma adicional de 500 000 dólares a cada uno de los demandantes que se arreglaron originalmente. Sabiendo que un arreglo mayor habrá de ser tan costoso para el demandado que será improbable que lo acepte, el quinto demandante experimentará una presión para aceptar algo mucho más cercano a 2 millones de dólares que a 2.5 millones, aunque sea el más optimista de los demandantes y crea que, si no fuera por la cláusula de la nación más favorecida, el demandado se arreglaría con él por 2.5 millones de dólares. Puede que se arregle con el demandado por 2.1 millones de dólares, lo que significa que el costo del arreglo para el demandado sería de 2.5 millones porque tendría que pagar otros 100 000 dólares a cada uno de los otros cuatro demandantes.

Empero, lo que socava al análisis es el hecho de que el demandado puede ser capaz de renegociar las cláusulas de la nación más favorecida con los demandantes que llegaron originalmente a un arreglo. El quinto demandante sabe que, si se rehúsa a acordar un arreglo, el acusado, temiendo el resultado de un juicio, podría acudir con los demandantes del arreglo original para decirles que, si esperan obtener más dinero gracias a un generoso arreglo acordado entre él y el quinto demandante, tendrán que poner dinero de su parte a fin de que así le alcance para cubrir ese magnánimo trato. A sabiendas de esa posibilidad, el quinto demandante podría aguantar más tiempo en espera de un arreglo más cuantioso, y la estrategia planeada podría desbaratarse.

[11] Véase Kathryn E. Spier, "The Use of 'Most-Favored-Nation' Clauses in Settlement of Litigation", 34 *RAND J. Econ.*, 78 (2003).

XXII.6. Cómo afectan al arreglo las reglas del procedimiento

Las reglas del procedimiento pueden afectar la tasa de arreglos. Consideremos primero las reglas que gobiernan el descubrimiento antes del juicio. Es probable que un intercambio completo de la información que poseen las partes facilite los arreglos al permitir que cada una de ellas se forme una estimación más correcta y, por ende, más convergente en general del resultado probable del caso; y el descubrimiento antes del juicio permite que cada parte obligue a su oponente a revelar la información relevante que posee. ¿Pero es necesaria la compulsión? El proceso de la discusión del acuerdo inducirá cierto intercambio de información aun en ausencia de todo derecho de descubrimiento previo al juicio.[12] Lo hará por la misma razón por la que los estudiantes aceptan que se revelen sus calificaciones a los posibles empleadores (véase antes I.4). En ausencia de toda información acerca de las pruebas favorables para él que están en posesión del demandado y que se presentarán en el juicio si el caso llega tan lejos, el demandante supondrá que tales pruebas —en un caso de discriminación en el empleo, podría tratarse de las evaluaciones del desempeño del demandante— son de utilidad promedio para él. Si en realidad tienen una utilidad por debajo del promedio para el demandante, el demandado las revelará voluntariamente al demandante a fin de inducir una demanda de arreglo menor. Si, por otra parte, el demandado no otorga voluntariamente ninguna información acerca de las pruebas que posee, el demandante inferirá que éstas tienen probablemente una utilidad por encima del promedio para él, y aumentará en consecuencia su demanda de arreglo. Si ni aun así se entregan las pruebas, el demandante elevará más aún su demanda, y este proceso continuará hasta que se entreguen o el demandante concluya que debe tratarse de pruebas de la máxima utilidad para él. El mismo estilo de razonamiento aplicado por el demandado a las pruebas que posea el demandante asegurará el ofrecimiento voluntario, por parte del demandante, de cualesquiera pruebas que sean útiles para el demandado, excepto las que sean de máxima utilidad para él.

Lo problemático de este enfoque, y una razón para que la revelación sea obligatoria en el contexto de la litigación pero no en el contexto comercial, es que un litigante no conoce todas las pruebas en posesión de su oponente: en nuestro ejemplo, el demandante sabe que el demandado tiene evaluaciones del desempeño, así como un empleador sabe que un solicitante de empleo tiene un registro de calificaciones de estudiante, pero no sabe si el demandado tiene también notas informales o memorandos acerca del demandante. Otra razón es que si fracasa una negociación comercial las partes seguirán

[12] Steven Shavell, "Sharing of Information Prior to Settlement or Litigation", 20 RAND *J. Econ.,* 183 (1989); Bruce L. Hay, "Civil Discovery: Its Effects and Optimal Scope", 23 *J. Legal Stud.,* 481 (1994).

caminos separados; si fracasa una negociación de arreglo, las partes se van a un juicio en el que la sorpresa tiene valor estratégico. Cada parte tiene un incentivo para retener información en la negociación del arreglo, sabiendo que si fracasan las negociaciones será la información más valiosa para él en el juicio si el oponente no ha tenido oportunidad de preparar una réplica para ella.

Si ninguna parte puede suponer por la ausencia de revelación que su oponente debe de tener pruebas muy favorables para ella, la falta de revelación no podrá inducir a esa parte a cambiar su demanda u oferta de arreglo, dado que no es posible hacer inferencia alguna a partir de la falta de revelaciones acerca de la información privada que posea el oponente. No todos los registros de calificaciones estudiantiles se revelarían si ello le costara 100 dólares al estudiante; y entonces el empleador tendría dificultad para obtener inferencias confiables acerca de la ausencia de revelación. El descubrimiento previo al juicio —que las pruebas del oponente se revelen por compulsión— podría así aumentar la probabilidad del arreglo al inducir al demandado a hacer una oferta mayor al demandante tras descubrir que su argumentación es más fuerte de lo que pensaba el demandado, pues el demandante ocultó la información por razones estratégicas. El aumento podría ser pequeño si el propio proceso del arreglo obtiene la mayor parte de la información privada de las partes, pero acabamos de ver razones por las cuales podría ocurrir algo diferente.

La Regla 35 de las Reglas Federales del Procedimiento Civil, una regla a la que se acogen con gran frecuencia los acusados en las acciones por lesión personal, autoriza a una de las partes a exigir que su oponente sea examinado por un médico designado por esa parte, si está en discusión la salud o la buena forma del oponente. Supongamos que el demandante se encuentra menos gravemente lesionado de lo que habría creído el demandado si no hubiese podido exigir un examen por el médico del demandado. Entonces éste no estará dispuesto a hacer una oferta de arreglo tan grande como la que haría antes del examen, cuando exageraba la medida de las lesiones del demandante; pero la oferta de arreglo mínima del demandante no se ve afectada, porque el examen no revela presumiblemente ninguna información nueva para él acerca de la medida de sus lesiones. Por lo tanto, se reduciría la probabilidad de un arreglo, a menos que, más allá de que el examen no le revele ninguna información nueva al demandante, sí lo convenza de que tendrá menos probabilidades de recuperar lo concerniente a daños sustanciales en el juicio de lo que había pensado, caso en el cual él rebajará su precio mínimo por el arreglo. En los casos donde el examen convenza al demandado de que las lesiones del demandante eran más graves de lo que creía el demandado, la Regla 35 aumenta la probabilidad del arreglo al empujar al acusado a incrementar su precio máximo por el arreglo, expandiendo así la gama habida en el acuerdo. De este modo, el efecto neto de la regla sobre el acuerdo no resulta claro. Además,

tampoco hay claridad en la cuestión de si, en el caso de ser abrogada la Regla 35, habría menos exámenes físicos practicados por médicos que designaran los acusados. La renuencia del demandante a ser auscultado por un médico elegido por el demandado sería (siempre y cuando el médico gozara de reputación e independencia) un indicio de que el demandante seguramente estaba exagerando acerca de su lesión… y el demandado podría sostener este aserto ante el jurado.

Un estudio de la administración judicial sostiene que si se considera el interés de un demandante ganador en el fallo desde la fecha del accidente (o de otro evento que origine su reclamación) no se afectará la tasa del arreglo, aunque el demandante tenga una tasa de descuento mayor que la del demandado.[13] De hecho, la adición de un interés anterior al fallo tenderá a reducir la probabilidad de un arreglo, cualesquiera que sean las tasas de descuento. Supongamos que antes de la adición del interés, y omitiendo los costos de la litigación y del arreglo, el valor esperado de la litigación para el demandante era de 120 dólares y la pérdida esperada para el demandado era de 100 (éste fue el ejemplo utilizado en el estudio citado). Si se añade el interés, digamos 6% anual, la ganancia esperada para el demandante aumentará a 127.20 y la pérdida esperada para el demandado a 106 dólares. La diferencia es mayor que sin el interés —21.20 en lugar de 20— y esto aumentará la probabilidad de la litigación al aumentar el miembro izquierdo de la desigualdad (1) más que el miembro derecho. Este resultado se aplica, aunque las partes tengan tasas de descuento diferentes.[14]

Así como el interés antes del fallo aumenta la probabilidad de la litigación al aumentar lo que está en juego, la demora pareciera reducir la probabilidad

[13] Hans Zeisel, Harry Kalven, Jr., y Bernard Buchholz, *Delay in the Court,* 133-136 (1959).

[14] Supongamos que, antes de la adición del interés, lo que está en juego son 10 000 dólares; la probabilidad subjetiva de ganar del demandante es de 80%; su tasa de descuento es de 15% anual, con una demora esperada de dos años; la probabilidad subjetiva de ganar del demandado es de 60% (es decir, el demandado cree que el demandante tiene una probabilidad de ganar de 40%); y su tasa de descuento es de 10%. Entonces la oferta de arreglo mínima del demandante (omitiendo los costos de la litigación y del arreglo) será de 6 049 dólares y la oferta máxima del demandado será de 3 306, una diferencia de 2 743. Si ahora aumentamos el fallo en 12% (dos años de intereses), las tres cifras se convierten en 6 775, 3 702 y 3 073 dólares, respectivamente. Si invertimos las tasas de descuento de las partes, entonces la oferta mínima del demandante, la oferta máxima del demandado y la diferencia, antes de la imposición del interés, son de 6 611, 3 024 y 3 587 dólares, respectivamente; después del interés, las cifras respectivas serían 7 405, 3 387 y 4 017. Invirtamos ahora las mejores ofertas del demandante y del demandado en nuestro ejemplo original, de modo que la oferta mínima del demandante sea de 100 dólares y la oferta máxima del demandado sea de 120. Nuestras conclusiones en lo concerniente a los efectos de añadir los intereses anteriores se revertirían. La oferta de arreglo del demandado se elevaría en mayor medida que la del demandante, con lo que se incrementaría el área de traslape entre ambas ofertas. Antes de la adición de interés, el acusado habría establecido un acuerdo a un precio entre 106 y 127.2 dólares, un intervalo más grande. Pero éste no es un ejemplo realista, puesto que, si la oferta de arreglo del demandado es más cuantiosa que la del demandante, para empezar, como se da por sentado en el ejemplo, el caso llevará a un acuerdo de cualquier modo.

de la litigación al disminuir lo que está en juego, suponiendo que las partes tienen tasas de descuento positivas y, por este motivo, valoran un dólar en su posesión hoy en día más que un dólar en el futuro. Pero esta conclusión debe calificarse en tres sentidos:

1. Si la tasa de descuento del demandado es mayor que la del demandante, la demora podría reducir la probabilidad de un arreglo al hacer que la oferta máxima de arreglo del demandado disminuya más deprisa que la diferencia existente entre las ofertas.
2. La demora aumenta la incertidumbre acerca del resultado porque la predicción es más difícil de hacer entre más prolongado sea el periodo futuro relevante; la demora en el establecimiento del arreglo exige que las partes consideren las probabilidades de que decaiga el valor de las pruebas, ocurran cambios del personal judicial, surjan nuevos precedentes, etc. Y, como hemos visto, puede esperarse que la incertidumbre disminuya las probabilidades de un arreglo.
3. Si todos los costos de litigación pueden diferirse hasta el juicio, tales costos disminuirán (a través del descuento) a la misma tasa que lo que está en juego, de modo que no cambiará la razón de las mejores ofertas de arreglo de las partes. Éste es el punto más débil. Pues no todos estos costos *pueden* diferirse hasta el juicio; y, aun si pudieran diferirse, el intervalo del arreglo se estrecharía aunque la razón de las ofertas fuese la misma. Por ejemplo, si las tasas de descuento de las partes son idénticas, el efecto de la demora consistiría en que las ofertas de arreglo de las partes se redujeran notablemente en una escala proporcional: de 120 000 y 100 000 dólares bajarían, respectivamente, a 80 000 y 75 000 dólares.

¿En qué punto, dentro del curso de una litigación, es probable que sobrevenga un arreglo? Según las apariencias, la probabilidad de un arreglo aumentaría conforme una demanda legal avanzara pasando por el descubrimiento previo al juicio y otras fases preparatorias hasta llegar al juicio. Las partes irían obteniendo más y más información acerca del resultado probable en el juicio y, por consiguiente, sus estimaciones de ese desenlace serían más y más convergentes. No obstante, esto pasa por alto el hecho de que, conforme progresa una demanda, el costo incrementado de la litigación *versus* el arreglo sufre una declinación. Así pues, por un lado, decaen los beneficios que se esperaban de una litigación (y éstos son una función del optimismo mutuo de las partes, sentimiento que debería decrecer a medida que éstas se enteran más acerca del caso), pero, por otro, ocurre lo mismo con los costos… si se hace caso omiso de los costos hundidos, como lo harán la mayoría de las personas racionales. De hecho, muchos casos se arreglan en vísperas del juicio; esto puede sugerir que el primer efecto es el que domina. También sugiere que

se puede establecer una analogía con el juego del gallina. Ninguna parte quiere echarse para atrás al ser el primero en sugerir un acuerdo.

Con todo, pocos son los casos que se resuelven con un acuerdo durante el estado de pendiente de una apelación, incluso después de que la argumentación oral puede haber dejado a las partes sobre aviso respecto a las inclinaciones de los jueces respecto a la apelación. El costo de una apelación para las partes es muy bajo, y luego de que el caso ha sido puesto por escrito y argumentado llega a cero.[15] Aun así, si los jueces mueven la mano hacia un lado, el intervalo o gama de la negociación se estrechará, lo que debiera incitar a obtener un arreglo acortando las negociaciones. Pero el abogado que sabe con un grado de virtual certeza que habrá de ganar se mostrará renuente a tener un arreglo, ya que desea que el triunfo quede registrado a fin de ayudarlo a atraer litigantes en el futuro. Y, debido a que el abogado que ha salido derrotado en el tribunal de instrucción no habrá de resarcirse de su pérdida obteniendo una resolución del caso en la apelación —por el contrario: un arreglo hará que se esfumen sus oportunidades de ser reivindicado mediante una revocación del juicio de la corte de menor escala por el tribunal de apelaciones—, será renuente a recomendarle a su cliente que se busque un acuerdo.

XXII.7. La decisión de llegar a un arreglo y la evolución del derecho común

Si la litigación es más probable entre mayor sea lo que está en juego, ¿podría explicar esto la tendencia aparente del derecho común, discutida en diferentes partes de este libro, a desarrollar normas de conducta eficientes?[16] Comparemos dos normas, una de las cuales conduce al doble del número de accidentes que la otra sin una reducción compensatoria del costo de las medidas de seguridad. La norma ineficiente conducirá a una litigación mayor que la norma eficiente, de modo que da a los tribunales mayor oportunidad para revisarla. Supongamos que, por mera casualidad, en el curso de esta revisión encuentran los tribunales el resultado eficiente. El número de accidentes disminuirá; la

[15] No es cero principalmente porque podría haber una posibilidad de apelación ante un tribunal superior y porque el tribunal de apelación podría ordenar nuevas actuaciones en el tribunal de instrucción.

[16] Una cuestión muy discutida. Véanse por ejemplo Paul H. Rubin, "Why is the Common Law Efficient?", 6 *J. Legal Stud.,* 51 (1977); George L. Priest, "The Common Law Process and the Selection of Efficient Rules", 6 *J. Legal Stud.,* 65 (1977); John Goodman, "An Economic Theory of the Evolution of the Common Law", 7 *J. Legal Stud.,* 393 (1978); William M. Landes y Richard A. Posner, "Adjudication as a Private Good", 8 *J. Legal Stud.,* 235, 259-284 (1979); Jack Hirshleifer, "Evolutionary Models in Economics and Law", 4 *Res. in Law & Econ.,* 167 (1982); Peter H. Aranson, "Models of Judicial Choice as Allocation and Distribution in Constitutional Law", 1990 *BYU. L. Rev.,* 745, 795-807 (1990); Martin J. Bailey y Paul H. Rubin, "A Positive Theory of Legal Change", 14 *Int'l Rev. Law & Econ.,* 467 (1994); Mark J. Roe, "Chaos and Evolution in Law and Economics", 109 *Harv. L. Rev.,* 641 (1996).

tasa de litigación disminuirá, y así será menos probable que los tribunales revisen la norma en el futuro. Por lo tanto, habrá a lo largo del tiempo una tendencia hacia la eliminación de las normas menos eficientes y su sustitución por normas eficientes, porque es menos probable que éstas sean revisadas y, por lo tanto, será menos probable que, una vez adoptadas, sean descartadas.

Hay varias objeciones a esta teoría de la evolución de la ley, empero. Una es que este tipo de proceso evolutivo a ciegas se tardaría mucho más tiempo en producir un conjunto eficiente de normas que la duración del periodo transcurrido desde el inicio del derecho común. Otra es que la distribución de los costos de las normas es más importante para la cantidad de litigación que la totalidad de tales costos. Una norma ineficiente cuyos costos sociales se dispersan ampliamente no provocará esfuerzos de litigación correctora porque ninguno de los perjudicados por ella tendrá un interés considerable en que se cambie la norma, mientras que por otra parte una norma eficiente que impone costos concentrados (digamos en una industria particular) bien podrá provocar tales esfuerzos. Además, una norma ineficiente podría reducir en lugar de aumentar la cantidad de litigación; por ejemplo, una que fijara un castigo excesivo para la conducta de los demandados potenciales. Por último, la teoría evolutiva omite el papel del precedente. Si una norma es ineficiente pero la *stare decisis* importa para el tribunal, el efecto de la litigación frecuente bajo la norma podría ser el de *solidificarla*. Y si la *stare decisis* compite con otros valores judiciales, se vuelve esencial la especificación de tales valores. Si se trata de valores favorables para la eficiencia, se acelerará la marcha hacia ésta, porque las normas ineficientes serán llevadas automáticamente una y otra vez ante los tribunales para su revisión. Pero, si son valores contrarios a la eficiencia, el efecto de la litigación más frecuente de normas ineficientes que de normas eficientes será una aceleración de la marcha del derecho alejándose de la eficiencia.

Un punto relacionado es que, cuando los demandantes pueden elegir en cuál tribunal se ha de entablar la demanda —por ejemplo, cuando demandan a una corporación que realiza negocios en varios estados—, se puede esperar que elijan un tribunal que sea favorable a ellos. Esto tenderá a asignarle un rumbo a favor de los demandantes a la evolución de la ley, una desviación no necesariamente congruente con la eficiencia, si bien se trata de una orientación menos pronunciada que la que se indujo con el método a través del cual se compensó a los jueces ingleses hasta las postrimerías del siglo XVIII. Se les pagaba con base en los derechos de presentación y, por ello, tenían un interés financiero en juego al fomentar la presentación de demandas en su tribunal, en especial dado que había varios tribunales en competencia entre los que los demandantes podían elegir dónde poner la demanda.[17]

[17] Daniel Klerman, "Jurisdictional Competition and the Evolution of the Common Law", 74 *U. Chi. L. Rev.,* 1179 (2007). Véase también Klerman, "Legal Fictions as Strategic Instruments" (USC Law School, 14 de agosto, 2009). Para un estudio empírico de una norma —la doctrina de la

XXII.8. La mediación como un auxiliar del arreglo

En el contexto del arreglo, la "mediación" se refiere al empleo de alguien neutral para ayudar a las partes a llegar a un acuerdo. Dado que el mediador (a diferencia de un arbitrajista) no está facultado para imponer un acuerdo, nos preguntamos cómo puede la mediación volver más probable un arreglo. La respuesta requiere considerar ciertos obstáculos que se interponen en el acuerdo:

1. Como antes vimos, por razones estratégicas las partes podrían resistirse a mostrar sus cartas en negociaciones de arreglo, lo que vuelve más difícil para cada parte determinar la solidez de la posición de la otra parte a fin de calcular el intervalo realista del arreglo.
2. También como se sugirió antes, cada parte puede resistirse a sugerir un arreglo por temor a denotar debilidad y hacer que la otra parte endurezca sus condiciones para el acuerdo.
3. Si las partes emplean el método convencional de la negociación paso a paso, en el que cada parte empieza con su posición "ideal" y gradualmente retrocede hasta llegar a la convergencia, cada parte se resistirá a alejarse más de uno o dos pasos de su posición inicial, de nuevo por el temor de señalar una debilidad. La otra parte elucubrará: "Puesto que mi oponente empezó en X, pero pronto se movió a X – 1 y luego a X – 2, me mostraré terco y veré hasta dónde está dispuesto a llegar". Si ambas partes piensan de este modo, la convergencia es improbable.
4. Los abogados tienen sus propios intereses pecuniarios y de otras clases que podrían empujarlos en una dirección opuesta a la que su cliente desearía seguir si supiera cuál es la situación. El abogado, especialmente el de una corporación grande, podría estar más ansioso que su cliente por llegar a un arreglo del caso antes de llegar a juicio, porque podría sentir mayor aversión por el riesgo. Puede tener relativamente pocos juicios, y una derrota dañará su reputación. (Esta tendencia de los abogados involucrados en los juicios a evitarlos puede ir aumentando conforme el número de juicios declina, según ocurre y tal como veremos en el siguiente capítulo.) Aun así, según lo advertimos anteriormente, los incentivos pueden apuntar hacia la dirección contraria cuando las negociaciones en pos del arreglo ocurren mientras un caso aguarda a que se tome una decisión por apelación (véase antes XXII.6).

El incentivo del cliente para lograr un arreglo puede también diferir del que tiene el abogado. Éste puede tener enfrente muchos juicios, lo que equi-

pérdida económica (véase antes VIII.2) en casos de construcción— que no ha evolucionado hacia la eficiencia, véase Anthony Niblett, Richard A Posner y Andrei Shleifer, "The Evolution of a Legal Rule", 39 *J. Legal Stud.,* 325 (2010).

vale a una cartera diversificada de casos, y puede asimismo albergar el deseo de desarrollar en beneficio de su compañía una reputación de no dar por fruto el arreglo de casos, con la finalidad de aumentar el riesgo y los gastos de los demandantes potenciales.

¿Cómo interviene la mediación en estas cuestiones?

1. Dado que el mediador puede reunirse con las partes por separado y sus discusiones con ellas son confidenciales, es probable que las partes sean más francas con él que entre sí, lo que le permitiría formarse una impresión más correcta que la de ellas mismas acerca de las ventajas y desventajas reales de sus respectivas posiciones, y comunicarles esta impresión de un modo creíble. Pero, para lograr esto, el mediador no debe ser sólo un conducto de información entre las partes opuestas, sino también un impedimento para la comunicación *transparente* entre ellas.[18] Cuando un mediador formula una propuesta a una de las partes, esa parte inferirá que la propuesta refleja información transmitida al mediador por la otra parte. Con todo, mientras la información sea alterada por el mediador, la otra parte estará entregando menos, en términos de una información estratégicamente valiosa (para el caso de que la mediación fracasara y el caso se vaya a juicio), que si debiera comunicarse cara a cara con su oponente.
2. Dado que el mediador puede hacer propuestas para un arreglo, ninguna de las partes necesita saber si una propuesta que se le formula emana de la parte contraria o del mediador. Por lo tanto, ninguna de las partes necesita identificarse como la que hizo la primera propuesta de arreglo, señalando así una debilidad. Las partes tienen así un incentivo mayor para hacer una propuesta al mediador para la transmisión (encubierta) a la otra parte que el incentivo que tendrían para hacer una propuesta directamente a la otra parte.
3. Las partes podrían hacer sus ofertas "ideales" al mediador, tratando de mostrarle fortaleza, como lo harían en una negociación sin intermediario. Pero el mediador no tiene que transmitir esa oferta a la otra parte, a menos que la parte oferente insista. Puede saltarse el método de negociación paso a paso reduciendo la escala de la propuesta hecha originalmente por una parte antes de presentársela a la otra. En esa forma, ninguna de las partes se ve inducida a señalar su fortaleza rehusándose a retroceder de una posición de negociación inicial extrema.
4. El mediador puede aliviar el problema de los costos de agencia incluyendo a los clientes en sus reuniones de mediación.

[18] Véase Jennifer Gerarda Brown e Ian Ayres, "Economic Rationales for Mediation", 80 *Va. L. Rev.,* 323 (1994). Véanse también Jeff Hawkins y Neil Steiner, "The Nash Equilibrium Meets BATNA: Game Theory's Varied Uses in ADR Contexts", 1 *Harv. Negotiation L. Rev.,* 249 (1996); Steven Shavell, "Alternative Dispute Resolution: An Economic Analysis", 24 *J. Legal Stud.,* 1 (1995).

He aquí dos hallazgos de un estudio empírico de la mediación:[19] es menos probable que la mediación logre llegar a un arreglo cuanto más grandes sean los intereses en juego en el caso; cuanto más dure una mediación, es más probable que tenga éxito. Tienen sentido económico estos hallazgos. Entre más grandes sean los intereses en juego, más probable será que los costos de litigación, por más que sean superiores al costo de mediación, no sean lo suficientemente grandes para persuadir a ambas partes de que se habrán de encontrar en una mejor situación si llegan a un arreglo que si litigan. No obstante, entre más tiempo dure la mediación, aun cuando esto signifique que el costo de litigación con respecto al de mediación irá en declive, mayores serán los indicios de que las partes opinen que la mediación tiene probabilidades de cosechar el éxito; de otro modo, la habrían suspendido en una fase más temprana en virtud de que no les habría parecido acertado incurrir en los costos adicionales.

XXII.9. Una revisión de las reglas de la responsabilidad

Un análisis completo de la elección entre reglas alternativas de la responsabilidad requiere considerar los costos procedimentales de cada una, lo que puede ilustrarse preguntándonos de nuevo si la eficiencia mejoraría remplazando la responsabilidad de la negligencia con la estricta (véase antes VI.5). Supongamos que el número de los accidentes permaneciera constante. Sin embargo, la adopción de la responsabilidad estricta conduciría a un aumento de las reclamaciones de perjuicios, porque sería mayor el alcance de la responsabilidad. Por lo tanto, si la fracción de las reclamaciones que se van a juicio permaneciera igual que bajo un estándar de negligencia, y si cada juicio fuese tan costoso bajo el nuevo estándar como bajo el antiguo, aumentarían los costos agregados de la litigación de cuasidelitos.

Los dos "si" son cuestionables. Dado que la responsabilidad estricta elimina un problema importante de la litigación de cuasidelitos, el del cuidado tomado por cada parte, se reduciría el monto de la incertidumbre implicada en el pronóstico del resultado de la litigación si la reclamación no se arregla; y una reducción de esa incertidumbre debiera reducir la fracción de las reclamaciones que se van a juicio. La simplificación de los problemas podría conducir también a una reducción del costo de cada juicio, aunque esto es incierto. La eliminación del problema de la negligencia tendería a aumentar el valor esperado de la reclamación del demandante, y esto podría llevarlo (aunque no necesariamente) a gastar más dinero en su caso antes que me-

[19] Douglas A. Henderson, "Mediation Success: An Empirical Analysis", 11 *Ohio St. J. on Disp. Resol.,* 105, 144 (1996). Adviértase la analogía con la subasta de las franquicias de televisión por cable para el participante que ofrece el mejor trato para los consumidores antes que el precio más alto por la franquicia. Véase antes XIV.11.

nos; pero el demandado podría gastar menos.[20] Empero, aun si el costo del juicio fuese menor con la responsabilidad estricta, el efecto sería, al estrechar la brecha existente entre los costos de la litigación y los del arreglo, el de volver a la litigación relativamente más atractiva que con el estándar de la negligencia, lo que conduciría a un aumento de la proporción de las reclamaciones que se van al juicio.[21]

Hasta ahora hemos supuesto que el número de los accidentes no cambia cuando se sustituye la negligencia por la responsabilidad estricta. Este supuesto podría ser injustificado. Una regla de responsabilidad estricta es más definida que una de negligencia y, por lo tanto, es probable que se administre con menos errores. El error jurídico reduce directamente la eficiencia del sistema de responsabilidad y, al aumentar el número de los accidentes, aumenta el número de las reclamaciones y, por ende, el gasto administrativo de un sistema de responsabilidad.

XXII.10. Negociación de la declaración de culpabilidad, la reforma del procedimiento penal y la tesis de la inutilidad

Contraparte que ofrece el procedimiento penal a las negociaciones del arreglo en el procedimiento civil, la negociación de la declaración de culpabilidad es objeto de críticas: se dice que le niega al demandado el derecho a las salvaguardas procedimentales de un juicio y que conduce a sentencias reducidas. Ninguna de estas críticas es persuasiva para un economista. Si un arreglo no mejorara la situación de ambas partes de un caso penal en relación con la que se tendría si se fueran al juicio, una o la otra invocaría su derecho a un juicio; por lo tanto, el acusado penal se ve compensado por renunciar a las salvaguardas procedimentales a las que tendría derecho en un juicio. Dado un presupuesto fijo para la persecución de los delitos, las sentencias típicas serán probablemente más severas, antes que más ligeras, si se permite la negociación de la declaración de culpabilidad. El fiscal podrá usar el dinero ahorrado en esas negociaciones (ya que, como ocurre en los arreglos civiles, una disposición de declaratoria de culpabilidad negociada resulta más barata que un juicio) para elaborar un argumento más sólido cuando la negociación fracasa; y, sabiendo esto, los acusados concederán mejores términos en la negociación. Esto implica que la negociación de la declaratoria de culpabilidad es peor

[20] Los determinantes de los gastos de litigación son complejos. Véase antes XXII.11.

[21] Si bien los costos del arreglo pueden ser menores también si el pronóstico del resultado de la litigación es más fácil porque los costos de dicho arreglo consisten en parte en los costos del tiempo tomado por los abogados para predecir el resultado de la litigación y, presumiblemente, esos costos son una función de la dificultad de la predicción, incluso una reducción proporcionalmente equitativa tanto de la litigación como de los costos por el arreglo se traducirían en una reducción en la diferencia absoluta entre ellos.

para los demandados penales en conjunto, aunque resulta sin duda mejor para los acusados individuales en la medida en que expande su conjunto de opciones a elegir. Y si viene a relajarse el supuesto de que existe un tope fijo para el presupuesto del fiscal, y si se da por sentado realistamente que los recursos a la disposición de los demandados penales son bastante limitados más allá del caso inusitado del delincuente provisto de caudales, entonces se fortalece la habilidad del fiscal para lanzar de manera creíble la amenaza de que la acusación seguirá si el demandado se rehúsa a la negociación de la declaratoria de culpabilidad.

Más sorprendente que el hecho de que los acusados en conjunto podrían encontrarse en mejor situación si se prohibiera la negociación de la declaratoria de culpabilidad negociada es el de que los *fiscales* podrían encontrarse en mejor situación. Supongamos que la negociación de la declaratoria de culpabilidad estuviera prohibida pero un acusado pudiera escoger declararse culpable en vez de someterse a un juicio y, si se declarara culpable, recibiera una sentencia más leve. Entonces la mayoría de los acusados culpables se declararían culpables y los fiscales se ahorrarían no sólo los costos del juicio sino también los costos de la negociación de la declaratoria de culpabilidad. El declararse culpable sin entrar en la negociación de la declaratoria de culpabilidad racionaliza el proceso "contractual" al consentir en un acuerdo mediante el que la oferta de indulgencia a cambio de una declaración de culpabilidad se convierte en una oferta de "tómalo o déjalo".

¿Habría de servir la medida de aumentar el número de jueces para reducir la negociación de la declaratoria de culpabilidad al incrementar la capacidad de los tribunales para conducir los juicios en vez de deshacerse de los casos a través de dicha negociación? No por necesidad. La negociación de la declaratoria de culpabilidad ocurre porque es un procedimiento más barato que la litigación para la resolución de las controversias. Por lo tanto, su incidencia se determina por los costos relativos de la negociación y de la litigación, y por el monto de la incertidumbre acerca del resultado de la litigación, factores que no se ven muy afectados por el número de los jueces (la incertidumbre podría ser en efecto mayor con más jueces). Aunque más jueces podrían permitir juicios más rápidos (véase después XXII.14), y los juicios más rápidos podrían afectar lo que está en juego para el acusado (y para el fiscal también[22]) y, por lo tanto, también los términos de la negociación, esto no debiera afectar la *cantidad* de las negociaciones.

[22] Los juicios celebrados con mayor rapidez podrían aumentar el costo de lanzarse a un proceso legal para la parte acusadora al requerir una expansión en el número de fiscales; ya no podría emplearse el procedimiento de hacer cola para aligerar la carga impuesta a la oficina del fiscal. Asimismo, en lo referente a los demandados que han sido encarcelados mientras aguardan el juicio, si éste transcurre en menos tiempo se reduciría el costo esperado del castigo y, de esta manera, el fiscal estaría dispuesto a ofrecerles un acuerdo atractivo para saltarse el proceso. Ambos puntos se apoyan en un supuesto sobre los incentivos del fiscal, a saber: que deriva una

Intuitivamente podría pensarse que proporcionarle asesoría a los acusados indigentes reduciría la proporción de contestaciones negociadas, pero la intuición es contraria a la teoría económica. El acusado que no tiene un defensor tendrá muy pocas perspectivas si elige ir a juicio, pero esto sólo significa que aceptará una sentencia negociada más larga que si tuviera la asistencia de un abogado. Si acaso, proporcionarle defensa al indigente debiera facilitar la negociación de declaratoria de culpabilidad, ya que un abogado defensor tiene mayores probabilidades que un acusado sin defensa de hacer una estimación correcta sobre el probable resultado de un juicio.

Si se prohibiera la negociación de la declaratoria de culpabilidad y no aumentara el número de los jueces —si, en otras palabras, aumentara varias veces la demanda de juicios penales sin que aumentara la oferta (a menos que los jueces dejaran de instruir casos civiles)—, el resultado sería o bien una caída en el número de acusaciones o bien un aumento enorme del periodo de espera de los juicios penales, puesto que en cierto momento una escasez de jueces debería conducir a colas más prolongadas a menos que se dictaran recortes drásticos en los honorarios del tribunal o se implementaran dramáticos cambios en la ley para reducir la demanda de juicios. (Probablemente podrían sobrevenir ambos efectos.) El costo esperado del castigo de las personas en libertad bajo fianza declinaría enormemente y el de las personas en prisión hasta el final del juicio aumentaría (a menos que pudieran argüir con éxito que su derecho constitucional a un juicio expedito había sido violado por la demora). La sentencia típica no debiera verse afectada por el hecho de que sea negociada o impuesta después de un juicio, pero la varianza de las sentencias aumentaría porque un juicio tenderá a conducir a una absolución o una condena más severa que la sentencia negociada. Una negociación es una solución intermedia entre la desiderata de las partes, en este caso el desiderátum del fiscal de imponer una sentencia más severa y la del acusado de ganarse una absolución. Desde la perspectiva del demandado penal que es indigente (es decir, alguien que no puede cubrir los honorarios de su abogado), no existe nada que lo induzca a la negociación a menos que la sentencia negociada sea más reducida de lo que podría esperar si se fuera a juicio.

El tiempo requerido para que los acusados penales a escala federal sean llevados a juicio ha sido abreviado por la Ley de Juicio Rápido, aunque con gran perturbación de los calendarios de juicios civiles de los jueces federales. ¿Vale la pena esa perturbación? La "línea" convencional sobre el juicio rápido es que la demora para llevar a un acusado a juicio es dura para el acusado al someterlo a una incertidumbre prolongada acerca de su destino, y es dura para la sociedad al reducir el costo esperado del castigo para quienquiera que tenga una tasa de descuento positiva. Pero estas dos afirmaciones no

utilidad positiva del número de condenas ponderadas por la duración de la sentencia. Véase después XXIV.3.

pueden ser ciertas para el mismo acusado; la demora mejorará o empeorará las cosas para él. Sin embargo, cada afirmación puede ser cierta para un grupo de acusados diferente: el de quienes están libres bajo fianza y el de quienes se ven obligados a permanecer en prisión en espera del juicio, respectivamente. Los juicios más rápidos aumentan los costos del castigo del primer grupo y reducen los del segundo grupo. Y para los acusados de este último (a quienes se niega la fianza) que son en efecto culpables, pero son absueltos o reciben una condena más corta que el periodo de su prisión anterior al juicio, ese encarcelamiento *es* su castigo, y toda medida que disminuya su duración reduce el costo efectivo del castigo… lo que hace precisamente el juicio rápido.

Desde la promulgación de la Ley de Reforma de la Libertad bajo Fianza de 1984, la detención anterior al juicio —es decir, la negativa a conceder la libertad caucional mientras se tramita el juicio, aun cuando pudieran depositar una fianza jugosa— se ha vuelto rutinaria en los juicios penales federales. El efecto de esta práctica controvertida, nocivo para los partidarios de las libertades civiles, es el de aumentar el costo del castigo esperado del acusado inocente y del acusado culpable, pero el primer efecto, el indeseable, se mitiga: *1)* por la Ley de Juicio Rápido, que minimiza el periodo de detención anterior al juicio, y *2)* por el hecho de que, en una época de elevadas tasas delictivas, pocos acusados absueltos son en realidad inocentes. Un fiscal que cuenta con un presupuesto fijo seleccionará los casos potenciales que pueden ganarse a un costo menor; ésos serán los casos donde los acusados son culpables, en realidad culpables a todas luces; pero, dada la pesada carga de la prueba para el fiscal, una parte de los acusados culpables a los que procesa serán absueltos.

Pudiera pensarse que el supuesto de que la mayoría de los acusados absueltos son culpables equivale a justificar no sólo la detención anterior al juicio sino también el abandono de *todas* las salvaguardas procedimentales tradicionales de los acusados penales. No es así. Lo que lleva al fiscal a seleccionar sus casos para detectar a los acusados que sean en efecto culpables es precisamente la dificultad que existe, dadas tales salvaguardas, para condenar a una persona inocente. La detención anterior al juicio es una práctica ominosa en las sociedades que carecen de nuestras elaboradas salvaguardas procedimentales en los juicios penales, o que gastan fuertemente en fiscales en relación con la cantidad de los delitos, o las dos cosas.

Los lineamientos para las sentencias federales, hasta que fueron degradados por la Corte Suprema en United States *vs.* Booker[23] al pasar de la condición de obligatoriedad a la de ser meramente consultivos, reducían notablemente el rango de discreción de los jueces federales. Esto puede haber reducido efectivamente los efectos de disuasión y prevención del castigo penal. Una gran discreción en las sentencias permite que el juez practique una forma de la discriminación de precios que consiste en decidir el castigo que resulta óptimo

[23] 543 U.S., 220 (2005).

en vista de las características particulares del acusado. Si el acusado parece pertenecer a una clase de personas que están fácilmente sujetas a la disuasión, una sentencia ligera podría bastar para disuadirlo a él, y a quienes sean como él, en el futuro; si el acusado es un delincuente endurecido e inveterado, podría requerirse una sentencia severa para este propósito, y, si no puede ser disuadido, incluso una sentencia severa lo mantendrá fuera de circulación, siendo incapaz de cometer delitos, por un largo tiempo. Si se promedian todas estas sentencias y se impone la misma sentencia a cada acusado, habrá una disuasión menor; la sentencia más severa se desperdiciará en los acusados fácilmente disuadibles, y la sentencia más ligera disuadirá insuficientemente (y dejará por demasiado tiempo en libertad) a los criminales endurecidos. Este problema puede remediarse "promediando hacia arriba"; pero los costos adicionales del encarcelamiento prolongado de los acusados fácilmente disuadibles se desperdician desde un punto de vista social.

Si bien por estas razones es de esperar que la uniformidad en las sentencias aminore la disuasión, puede incrementar la disuasión de los delincuentes que aman el riesgo al reducir la varianza de las sentencias (a ellos *les gusta* la varianza, pues conciben la duración de la sentencia como un juego de apuestas) y, por la misma razón, puede aminorar la disuasión experimentada por los delincuentes adversos al riesgo (a ellos no les gusta probar su suerte y arriesgarse a que les den una sentencia impredeciblemente prolongada en el caso de que resultaran encontrarse frente a un juez con la mano pesada en materia de delitos).

La uniformidad en las sentencias, tal como se hacía cumplir de acuerdo con los lineamientos antes de que éstos pasaran de ser obligatorios a tener un carácter consultivo, impidió los siguientes abusos por parte de la fiscalía: el de inducir a un acusado inocente a declararse culpable una vez que se le ofrece una sentencia nominal,[24] y el de amenazar a un acusado contra quien el gobierno ha elaborado una argumentación débil con que se le pedirá al juez que le imponga una sentencia en extremo severa (digamos, la sentencia estatutaria máxima, que es probable que sea considerablemente más elevada que la ubicada en el punto más alto del rango de sentencias de los lineamientos para el delito del acusado).

Es incierto saber hasta qué punto la creación de salvaguardas procedimentales elaboradas para los acusados penales puede hacer mucho a largo plazo para proteger al inocente, y específicamente si la creación de nuevos derechos procedimentales para los acusados penales por la Corte Warren en el decenio de 1960 hizo mucho. Los legisladores federales y estatales respondieron de diversas maneras. Una de ellas consistió en ampliar las sentencias estatutarias máximas, con el resultado de que, incluso si quedaban tras las rejas

[24] Véase Oren Bar-Gill y Oren Gazal Ayal, "Plea Bargains Only for the Guilty", 49 *J. Law & Econ.,* 353 (2006).

menos personas inocentes, las que quedaban convictas probablemente habían de cumplir sentencias más dilatadas. Otras respuestas legislativas a los nuevos derechos procedimentales fueron la de expandir la detención previa al juicio, la de autorizar un uso mayor de las grabaciones y otros instrumentos de vigilancia electrónicos, la de reducir el derecho de libertad provisional bajo caución y así aumentar la detención previa al juicio, la de crear nuevos delitos (dando así a los fiscales más opciones a la hora de acusar a los demandados penales), la de podar el *habeas corpus* y la de destinar más dinero para prisiones, policías y fiscales, así como menos para la defensa de los acusados penales en indigencia. El resultado fue un juego —en términos de la teoría de juegos— entre los tribunales, que son primordialmente responsables de la creación, mediante la interpretación constitucional, de nuevos derechos para los acusados penales, y las legislaturas.

Para formalizar esta discusión, sea que el costo esperado del castigo, una medida de disuasión, se denote por $CE = pS$, siendo p la probabilidad de la aprehensión y la condena, y S la sentencia. Si un derecho creado por los tribunales conduce a una reducción de p para los acusados inocentes y culpables (y ésa es la consecuencia más probable, porque un derecho que vuelve más difícil la condena de una persona inocente volverá más difícil también la condena de una persona culpable), y la legislatura desea mantener CE a su nivel anterior, podrá lograrlo aumentando S mediante una ley que aumente los castigos de los delitos o aumentando p mediante una reducción de los fondos destinados a la defensa de los acusados indigentes. Ambas medidas han sido en efecto respuestas legislativas, en los Estados Unidos, para lo que se percibía como excesos judiciales en la protección de los derechos de los acusados penales.

Adviértase la complicación creada por las respuestas legislativas para los tribunales que tenían la vista puesta en relajar las salvaguardas procedimentales impuestas por decisiones anteriores. Si a fin de compensar las salvaguardas que reducían los costos esperados del castigo la legislatura había incrementado la severidad del castigo, un relajamiento de las salvaguardas que incrementara la probabilidad del castigo (una p más alta en la fórmula del párrafo precedente) aumentaría los costos esperados del castigo (CE) por encima del nivel que pretendía alcanzar la legislatura.

XXII.11. Los gastos en litigación y la búsqueda del procedimiento eficiente

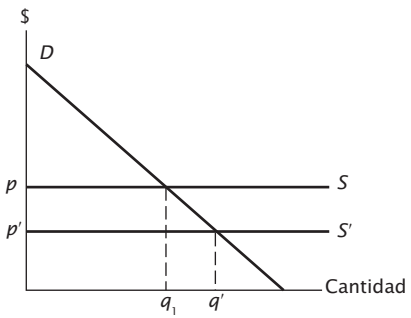
Una parte optimiza sus gastos de litigación hasta el punto en que un dólar gastado aumente el valor esperado de la litigación para él (al aumentar sus probabilidades de ganar) precisamente en un dólar. Pero toda decisión de gasto de una de las partes afecta la decisión del gasto de la otra, al alterar la probabilidad y, por ende, el valor esperado de un resultado favorable para la otra, en

una forma muy similar a lo que ocurre cuando todo cambio de la producción de un oligopolista altera la producción óptima de sus rivales (véase antes X.5). Si cada parte, al decidir cuánto gastará en la demanda, toma en cuenta el efecto de sus gastos sobre los gastos de la otra parte, entonces, como en el caso del oligopolio, no habrá un nivel determinado en el que el analista pueda decir que ninguna de las partes (si son racionales) tendría algún incentivo para hacer nuevos cambios en sus gastos.

Con todo, dado que estos gastos se compensan en gran medida, las partes de la litigación consideran a menudo mutuamente ventajoso ponerse de acuerdo para no incurrir en un gasto particular de litigación (por ejemplo, estipulando un hecho de tal manera que no tenga que establecerse por testimonio). Esto parecería ser algo bueno. Pero los gastos compensatorios en la litigación no son necesariamente dispendiosos desde un punto de vista social, al revés de lo que ocurre desde un punto de vista privado, ya que (dentro de ciertos límites: en cierto punto, los costos derivados de absorber más información superan a los beneficios) aumentan la probabilidad de una decisión correcta al proporcionar más información al tribunal.

Aun sin poder modelar precisamente las funciones de reacción de cada parte a las decisiones de gasto de la otra, podemos tener bases razonables para creer (de acuerdo con XXII.5, antes) que las partes gastarán probablemente más en la litigación entre mayor sea lo que está en juego. Por lo tanto, esperaríamos que los casos más grandes se decidieran correctamente en una proporción del tiempo mayor que la de los casos más pequeños.

Muchas reglas procedimentales y probatorias (en el siguiente capítulo examinamos sistemáticamente estas últimas) pueden considerarse diseñadas para aumentar la productividad de los gastos de litigación de las partes. Tenemos un ejemplo en la regla que permite que el juez tome nota judicial de los hechos obviamente ciertos, de modo que la parte a la que le corresponde la carga de la prueba no tenga que establecer el hecho con sus evidencias. El efecto de esta regla sobre los gastos de litigación de las partes se muestra en la gráfica XXII.2. *D* es el valor medio, para la parte, de diversas cantidades de



GRÁFICA XXII.2

pruebas, y *S* es el costo medio (igual al costo marginal) de esas pruebas. Una regla procesal (como la de la nota judicial) que disminuya el costo de las pruebas sin reducir su valor desplaza *S* hacia abajo, hasta *S'*, haciendo que el litigante aumente su compra de pruebas de *q* a *q'*. Sin embargo, el aumento de su gasto total (precio multiplicado por cantidad) dependerá de la elasticidad-precio de la demanda en la región de la curva de demanda entre *q* y *q'*. Si la elasticidad es menor a 1, el gasto total disminuirá; si es igual a 1, permanecerá constante; y si es mayor a 1, el gasto aumentará.

Un importante conjunto de reglas procedimentales se refiere a la reunión de las reclamaciones y de las partes (demandantes o demandados, o ambos) mediante instrumentos tales como las demandas colectivas (discutidas en la sección siguiente), la reunión permisiva y obligatoria, y las reclamaciones contrarias (contrademandas) permisivas y obligatorias. Consideremos la contrademanda obligatoria.[25] Si el demandado tiene una reclamación contra el demandante que surge del mismo hecho o la misma transacción que originaron la reclamación del demandante contra el demandado, este último debe promover esa reclamación en el juicio del demandante, en lugar de constituirla en la base de un juicio separado. La meta es la economización de los gastos de litigación y del tiempo de los tribunales. Podría parecer desconcertante que deba *obligarse* a las partes de un juicio a adoptar una medida que ahorrará costos para ambas. La explicación es que los costos de negociación tienden a ser elevados debido a la multitud de factores imponderables implicados, como las ventajas que el demandado podría percibir al ser capaz de elegir el foro en el que litigará su reclamación. Por lo tanto, el hecho de volver obligatorias ciertas reclamaciones contrarias ayuda a los demandantes; el *permitir* las contrademandas (reclamaciones vinculadas a la reclamación del demandante, pero sólo de manera vaga) ayuda a los demandados, obviamente, pero también a los demandantes en la medida en que resulta más barato para ellos litigar las dos reclamaciones en un solo caso que en dos. ¿Y acaso el rasgo de la obligatoriedad aumenta o disminuye la tasa de arreglos? La respuesta sugerida por el análisis económico es que así se reduce la tasa de arreglos, porque la reclamación del demandado aumenta lo que está en juego en el caso, pero no aumenta los costos de la litigación en la misma medida (éste es el aspecto economizador de la regla).

En este ejemplo, una medida destinada a reducir los costos de la litigación podría aumentarlos al aumentar la tasa de litigación. Éste es un problema general de la eficiencia procedimental. Consideremos el procedimiento muy elogiado del juicio "bifurcado". La cuestión de la responsabilidad del demandado se instruye por separado, antes de que se instruya la cuestión del remedio (por ejemplo, el monto de los daños), según la teoría de que si el demandado

[25] Véase William M. Landes, "Counterclaims: An Economic Analysis", 14 *Int'l Rev. Law & Econ.,* 235 (1994).

gana en la etapa de la litigación se evitará el costo de la litigación del remedio. Pero, al reducir el costo de los juicios sin reducir el costo de los arreglos, es probable que el procedimiento bifurcado aumente el número de los juicios, tanto al reducir la tasa de los arreglos como al volver rentables algunas demandas que no lo serían si sus costos fuesen mayores. El aumento del número de juicios podría incrementar los costos agregados de la litigación, aunque la litigación típica sea más barata porque en muchos casos no habrá la etapa de los daños.[26]

Y es posible que la litigación típica no sea más barata con la bifurcación; podría ser más cara porque el arreglo pudiera ocurrir más tarde. Supongamos que el demandado ignora si el demandante ha sufrido daños cuantiosos o escasos. El demandado formula entonces una oferta de arreglo que un demandante con daños escasos aceptaría, pero un demandante con daños cuantiosos rechazaría. El demandante con daños escasos rechazará la oferta y se irá a juicio, porque si gana en el juicio el demandado procurará entonces llegar a un arreglo, pensando que el demandante debe de haber sufrido daños cuantiosos. Sabiendo que así es como piensan los demandantes de daños escasos, el demandado no tendrá ningún incentivo para hacer la oferta inicial.[27]

Se afirma a menudo que los individuos ricos o las grandes empresas podrían tratar de abrumar a sus contrarios en la litigación mediante grandes gastos. Nuestra discusión del precio depredador (véase antes X.9) es relevante para evaluar esta preocupación. Si una de las partes prevé una sucesión de juicios legales similares, podría resultarle racional formular, en el primer juicio, una amenaza de abrumar a su contrario, y el cumplimiento de esa amenaza mediante un gran gasto en la litigación a fin de fortalecer la credibilidad de las amenazas contra contrarios subsecuentes. Así, sabiendo que podría ser racional que la parte cumpla la amenaza, el oponente podría rendirse.

¿Deberían sancionarse por lo tanto los gastos de litigación "depredadores"? Es probable que no. Como vimos en la discusión del problema paralelo en la legislación antimonopólica, resulta muy difícil saber si un gasto (un recorte en el precio es una forma de gasto) es "depredador" en algún sentido que connote una situación ingrata. Los gastos *excesivos* en la litigación son comunes en virtud de que los abogados son agentes imperfectos de sus clientes; también lo son los gastos compensatorios, que se asemejan a los gastos compensatorios en publicidad en que incurren los vendedores en competencia; pero el gasto dispendioso no equivale a un gasto depredador. En la práctica resulta también difícil distinguir entre los gastos de litigación depredadores y los grandes gastos en la litigación que sí tienen sentido, independientemente de cualquier motivación estratégica, simplemente porque una parte que prevé una

[26] William M. Landes, "Sequential *versus* Unitary Trials: An Economic Analysis", 22 *J. Legal Stud.*, 99 (1993).

[27] Véase Douglas G. Baird, Robert H. Gertner y Randal C. Picker, *Game Theory and the Law,* 256-259 (1994).

sucesión de juicios similares afronta mayores pérdidas en el primer juicio ya que su decisión podría tener un efecto, a través de la *stare decisis* o la *res judicata* (cosa juzgada), sobre el resultado de los casos subsecuentes.

En efecto, aun si de nuevo se excluye la necesidad de plantear consideraciones estratégicas, es posible que los gastos combinados de las partes *superen* lo que está en juego en su juicio sencillamente debido al carácter secuencial de los gastos. Supongamos que *J* es un millón de dólares para cada una de las partes, A y B; cada parte estima inicialmente que gastará 300 000 dólares en gastos legales; y cada una piensa que el gasto le dará una probabilidad de ganar de 60%. (Omitiendo el arreglo.) Después de que cada parte ha gastado 250 000 dólares, A decide que un gasto adicional suyo de 100 000 aumentará su probabilidad de triunfo de 60 a 75%. El valor esperado de este incremento es de 150 000 dólares, de modo que el gasto es rentable. B se preocupa; cree que el gasto de A disminuirá su propia probabilidad de triunfo de 60 a 45%, y contesta con su propio gasto de 150 000 destinado a anular el gasto de A. Y en efecto A revisa ahora sus probabilidades de triunfo y nota que sólo son de 60%. Esperando que B no responderá a un incremento menor de sus gastos, A gasta ahora otros 100 000 en peritos, instrumentos paralegales, descubrimiento de terceros, expertos en la selección de jurados, etc. A está equivocado, y B responde con sus propios 100 000 en gastos adicionales. Para este momento las partes han gastado conjuntamente, en litigación, una suma igual a *J*. Y podrían seguir adelante. O quizá no. Cada parte podría prever las reacciones de la otra varias movidas adelante, o podrían convenir ambas en limitar sus gastos. Pero si a las partes les resulta fácil entender las evaluaciones del contrario o negociar arreglos mutuamente satisfactorios, habrían arreglado el caso. Por lo tanto, el patrón esbozado antes, aunque patológico, podría no estar totalmente desprovisto de realismo.

XXII.12. El acceso a los remedios legales: honorarios contingentes, demandas colectivas e indemnización de los honorarios legales

Supongamos que un demandante tiene una reclamación de 100 000 dólares y una probabilidad de 50% de hacerla valer si cuenta con un buen abogado. El valor esperado de la reclamación es de 50 000 dólares y ello justificaría su gasto de hasta esa cantidad en honorarios del abogado para proteger el activo. (En este ejemplo suponemos que es neutral al riesgo.) Pero supongamos que la reclamación es su único activo. Esto no sería ningún problema de ordinario; podemos pedir prestada una suma sustancial con un activo como aval. Pero no siempre se puede pedir prestado con el aval de una reclamación legal. Los bancos y otras instituciones de préstamos podrían sentir aversión por el riesgo a causa de la regulación gubernamental de las instituciones financieras o podría resultarles costosa la estimación de la probabilidad de que la reclamación

pueda establecerse en el tribunal. Estos factores podrían volver prohibitivamente elevada la tasa de interés. Y muchas reclamaciones legales (notablemente la mayoría de las referidas a cuasidelitos) son legalmente intransferibles —a fin, se supone, de impedir el fomento de la litigación—, de modo que no valen nada como aval. Una posible justificación económica para esta regla es que, una vez que la reclamación es transferida, el demandante original, quien puede ser el testigo clave, tendrá menos incentivos para rendir testimonio de modo efectivo. Ello puede explicar por qué tales transferencias son raras, pero no por qué están prohibidas. Las reclamaciones por contrato son usualmente transferibles, lo que no es congruente con la regla de no transferencia para las reclamaciones sobre cuasidelitos.

En los estados se observa hoy en día una tendencia a relajar la proscripción tradicional de la venta o el financiamiento bancario de las reclamaciones legales. Los troles de patentes (examinados en el capítulo XI) ocupan la delantera. Muchos de ellos compran patentes. Su propósito al comprarlas es hacer una demanda por infracción. Esencialmente, están comprando el juicio de un propietario de una patente. Un atractivo de la venta de reclamaciones legales (o de la defensa de éstas) consiste en que disemina el riesgo. Supongamos que una compañía es la parte demandada en un juicio prolongado. Podría tener el deseo de vender la defensa del juicio a una compañía legal que recibiría unos honorarios sustanciales si la defensa prevaleciera, pero llevaría sobre sus espaldas el costo entero de la defensa si fracasara. Eso le ahorraría al abogado general a cargo de la defensa del demandado la pena de tener que informarle a su director ejecutivo que la defensa no solamente había fracasado, sino que le había causado a la compañía unos costos legales en verdad onerosos.

La solución tradicional para este problema de liquidez que afronta el demandante es el contrato de honorarios contingentes. El abogado presta sus servicios por una parte de la reclamación. El riesgo disminuye porque el abogado especializado en cuestiones de honorarios contingentes puede combinar muchas reclamaciones y minimizar así la varianza de los rendimientos. La especialización le permite también estimar los riesgos con mayor precisión de lo que podría hacerlo un prestamista convencional; y hay economías en el hecho de hacer que la misma persona o empresa evalúe el riesgo y lo asegure.

Se ha sostenido que los honorarios contingentes son a menudo exorbitantes. Pero aquí es fácil extraviarse. Un honorario contingente debe ser mayor que un honorario por los mismos servicios legales que se pagan a medida que se prestan o poco después. Un honorario contingente compensa al abogado no sólo por los servicios legales que presta sino también por el préstamo de tales servicios. La tasa de interés implícita sobre tal préstamo es elevada porque el riesgo del impago (la pérdida del caso, que cancela la deuda del cliente con el abogado) es mucho mayor que el de los préstamos convencionales, y el monto total de los intereses es mayor no sólo porque la tasa de interés sea

elevada sino también porque el préstamo puede quedarse insoluto durante muchos años, y sin pagos parciales periódicos, lo que es un instrumento para la disminución del riesgo soportado por el prestamista ordinario.

Pero si lo que está haciendo el abogado es un préstamo riesgoso de sus servicios, ¿por qué no debería ser el honorario contingente un costo de oportunidad de tales servicios con intereses, antes que una fracción del fallo o el arreglo? Hay tres respuestas económicas relacionadas. La primera es que el gasto óptimo en servicios legales depende de lo que esté en juego. Entre más sea lo que esté en juego, más gastará el demandado para demorar un fallo y, por lo tanto, será más lo que tenga que gastar el abogado del demandante en términos de tiempo y esfuerzo para ganar. Así, el costo de sus servicios será probablemente proporcional a lo que esté en juego (si bien irá en mengua a causa del componente de costos fijos que conllevan los costos de litigación), de modo que podrá expresarse como un porcentaje (si bien cabe más hablar de un porcentaje en declinación) de esa suma. Segundo, el hecho de hacer que los honorarios del abogado varíen con el éxito de su esfuerzo es un procedimiento para otorgarle un incentivo a fin de que haga un buen trabajo. Y tercero, se traslada más riesgo del demandante a su abogado. Dado que el abogado gana más cuando el demandante gana mucho, y menos cuando gana poco, disminuye la varianza de la ganancia esperada del juicio por el demandante tras la deducción de los honorarios del abogado.[28] Bien puede ser el abogado quien sobrelleve de modo superior el riesgo debido a que puede diseminar el riesgo de la litigación a través de una cartera de casos.

Un problema del honorario contingente es que, en cualquier situación de propiedad conjunta (y el contrato de honorarios contingentes convierte al abogado en efecto en un tenedor conjunto de la propiedad representada por la reclamación del demandante), cada propietario, como vimos en el capítulo III, podría carecer de un incentivo adecuado para explotar el derecho, porque una parte de la ganancia de sus esfuerzos en tal sentido va a dar a las manos de otra persona. Supongamos que se ofrece al abogado del demandante un arreglo de 100 000 dólares; si se va a juicio, hay una probabilidad de 90% de que el demandante gane 150 000, pero la instrucción del caso le costará al abogado 25 000 dólares de su tiempo; las partes sienten aversión por el riesgo; y el honorario contingente es de 30%. Si el demandante acepta el arreglo, obtendrá 70 000 dólares en términos netos, y el abogado recibirá 30 000. Si el

[28] Por ejemplo, supongamos que el intervalo esperado del juicio vaya de 0 a 100 000 dólares, donde cada punto en el intervalo corresponde a un ingreso igualmente probable, y el cobro contingente constituye 25% y, por consiguiente, la tarifa fija es equivalente a 12 500 dólares. (El juicio esperado suma 50 000 dólares, y el 25% de esa cifra son 12 500. Pero éste no es *realmente* el equivalente de la tarifa fija, puesto que con una tarifa de esta clase el demandante está asumiendo cabalmente el riesgo de la pérdida. La verdadera tarifa fija sería más baja porque no contendría un componente que compensara al abogado por asumir el riesgo.) Así pues, la gama de resultados esperados netos para el demandante oscila entre –12 500 y 87 500 dólares con el enfoque de la tarifa fija y entre 0 y 75 000 dólares con el enfoque de los honorarios contingentes.

caso se va a juicio, la ganancia esperada neta para el demandante aumenta a 94 500 [0.9 × (150 000 – 45 000)],[29] pero la ganancia esperada neta del abogado baja a 15 500 (45 000 × 0.9 – 25 000). Por lo tanto, hay aquí un conflicto de intereses entre las partes debido sólo al hecho de que al abogado no se le permite obtener la ganancia total de un juicio (la ganancia esperada neta del juicio es de (50 000 × 0.9) – 25 000, de modo que es positiva).[30] El conflicto se podría eliminar permitiendo la venta directa de todas las reclamaciones legales; las reclamaciones por contrato son ya transferibles.

Contra el contrato de honorarios contingentes y, de forma más clara, contra el enfoque de la venta directa puede argüirse que estos instrumentos fomentan más litigación. No por fuerza, aun cuando ciertamente facilitan que una persona sin liquidez o que sienta aversión por el riesgo promueva un juicio. Entre más probable sea un juicio si existe una violación de la ley que causa perjuicio, mayor será el efecto disuasivo de cualquier principio legal que se haga valer en el juicio y, por ende, será menos probable que los demandados potenciales observen la conducta prohibida. Más aún, un contrato de honorarios contingentes le da al abogado un incentivo mayor para declinar un caso débil que si se le pagara por horas,[31] porque el costo de la pérdida se traslada del cliente al abogado. Se trata de un instrumento de filtración.

Las reclamaciones muy pequeñas no crearían ningún problema para el sistema legal si no fuera por el componente fijo de los costos de la litigación. La gente invertiría simplemente menos cuando lo que estuviera en juego fuese menor. Pero, si el costo fijo puede repartirse entre muchas reclamaciones, podrían hacerse valer más reclamaciones, con una disminución consiguiente de los costos del error del sistema legal y, aun así, sin incurrir en costos de litigio prohibitivos. Desde hace largo tiempo han existido técnicas para reunir varias reclamaciones pequeñas en una sola reclamación suficientemente grande para justificar los costos del juicio… o, dicho de otro modo, para obtener economías de escala en la litigación. Una tienda de departamentos realiza esta función respecto de las reclamaciones de sus clientes contra los fabricantes cuyos productos vende. El cliente que compra mercancías defectuosas podría no tener en juego lo suficiente para demandar al fabricante, pero no vacilará en quejarse en la tienda de departamentos, la que remplazará la mercancía o devolverá el dinero al cliente, y, si varios clientes se quejan, reunirá estas reclamaciones y las presentará al fabricante. Si este último no está dispuesto a rem-

[29] 45 000 dólares es el cobro de 30% del abogado sobre un fallo de 150 000.

[30] Véanse en George B. Shepherd, "An Empirical Study of the Economics of Pretrial Discovery", 19 *Int'l Rev. Law & Econ.,* 245, 258 (1999), algunas pruebas de que, como lo pronostica el análisis económico, los abogados pagados en una forma contingente hacen menos descubrimientos antes de la instrucción que los abogados pagados por hora.

[31] Thomas J. Miceli, "Do Contingent Fees Promote Excessive Litigation?", 23 *J. Legal Stud.,* 211 (1994). Véase también Eric Helland y Alexander Tabarrok, "Contingency Fees, Settlement Delay, and Low-Quality Litigation: Empirical Evidence from Two Databases", 19 *J. Law, Econ. & Organization,* 517 (2003).

bolsar a la tienda sus costos para responder a las quejas de los clientes, la tienda podrá hacer una amenaza creíble de demandar al fabricante.

La acción de clase moderna generaliza esta técnica. Supongamos que los fabricantes de cepillos de dientes han conspirado para cobrar un precio monopólico. Millones de consumidores resultan perjudicados; el costo agregado podría ser sustancial; y sin embargo el perjuicio para cada consumidor podría ser sólo de unos cuantos centavos. Si se agregan todas estas reclamaciones en una acción de clase, lo que está en juego en la acción será suficientemente grande para sufragar los costos del juicio.[32]

El instrumento de la demanda colectiva podría parecer de escasa utilidad, no obstante, precisamente en el caso en el que más se necesita: cuando la reclamación individual es muy pequeña. Puede obligarse al demandado a pagar un fallo igual a los costos de su infracción, ¿pero a quién? Los costos derivados de identificar a los miembros del grupo y de dar a cada uno de ellos sus daños individuales (unos cuantos centavos, en nuestro ejemplo) podrían superar el monto de la condena. Es cierto que lo más importante desde un punto de vista económico es que el infractor se vea confrontado por los costos de su infracción —esto preserva el efecto disuasorio de la litigación—, no que les pague a sus víctimas. Y nuestro hincapié anterior en la importancia de la compensación para la parte perjudicada a fin de motivarla a operar la maquinaria legal e impedirle que tome precauciones excesivas (véase antes VI.4) no es aplicable aquí; lo que está en juego es demasiado pequeño para inducir a cualquier víctima a soportar cualquier carga de la obtención de un remedio legal. Además, la ausencia de un cliente "real" —no hay un reclamante que tenga un interés significativo en el fallo— hace que el abogado ocupado en dicho grupo se halle bajo una presión nimia para emplearse a fondo a fin de maximizar el beneficio recuperado por éste. Sus ganancias del juicio se determinan por el honorario legal que recibe y no por el monto de la condena. Nadie tiene un interés económico del tamaño de la condena, excepto el demandado, que tiene un interés en minimizarlo. El abogado del grupo se verá, pues, tentado a ofrecer un arreglo con el demandado por una condena pequeña (pequeña, esto es, con respecto a lo que una representación más vigorosa podría producir) y un honorario legal grande, sabiendo que tal oferta será atractiva para el demandado siempre que la suma de las dos cifras sea menor a la pérdida neta esperada por el demandado en caso de irse a juicio. (No es de sorprender que los juicios de instrucción en las demandas de acción de clase sean en extremo raros.)

Aunque el juez debe aprobar el arreglo de una acción de clase, los abogados controlan en gran medida su acceso a la información —acerca de los

[32] Si el fabricante no vende directamente al consumidor, no se permitirá que el consumidor demande, en una acción de clase o de otro modo; sólo los intermediarios tendrán el derecho de demandar. Ésta es una alternativa económica para la acción de clase del consumidor por las razones discutidas antes en X.14.

méritos de la reclamación, la cantidad de trabajo realizado por los abogados en nombre de la clase, los daños probables si el caso se va a juicio, etc.— que es vital para determinar lo razonable del arreglo. De ahí el peligro de lo que se llama "subasta en reversa". Un demandado enfrentado a la amenaza de múltiples demandas colectivas (como es común cuando una gran cantidad de personas tienen una reclamación potencial) podría buscar al consejero de grupo que tenga menores probabilidades de obtener una condena sustancial en caso de haber un juicio y ofrecerle un trato congruente en grandes honorarios para los abogados, pero sólo una indemnización moderada para el grupo. Las partes luego se unirán al pedirle al tribunal que apruebe el arreglo como un "arreglo global" de todas las demandas colectivas (y que dicte la continua prosecución de todas las demás demandas colectivas a fin de hacer efectivo el arreglo global), señalando las malas perspectivas de la litigación del consejero del grupo como una razón para que el tribunal apruebe el arreglo a pesar de su moderación.

Ya que ni el demandante ni el demandado están motivados para oponerse a que el consejero del grupo reciba una retribución excesiva, la responsabilidad de proteger al grupo recae en la corte del tribunal y es difícil de cumplir de forma adecuada; una dificultad especial radica en que los abogados que se especializan en litigios de acciones de clase pueden contar con pocos clientes que los remuneren o incluso con ninguno en absoluto, lo que es un obstáculo para determinar el valor de mercado de sus servicios. No obstante, la competencia puede ayudar a la corte a fijar una cuota razonable. Primero, los que sirven de abogados a los miembros del grupo objetor tienen un incentivo para pronunciarse contra el estipendio propuesto por el demandante y el demandado porque recibirán un pago de compensación si sus objeciones son sostenidas. Segundo, los miembros del grupo tienen permiso para optar por salirse de él (si éste busca reparación de daños), a levantar ya sea una demanda individual o una demanda colectiva por separado, o asimismo a poner objeciones al arreglo de grupo propuesto, incluida la propuesta de honorarios para los abogados. El objetor cuyas objeciones persuadan a la corte recibirá una recompensa en efectivo.

Por añadidura, los tribunales han comenzado a subastar la designación de consejero de grupo al abogado o la compañía legal que ofrezca el mejor trato para el grupo, es decir, los mayores beneficios restados de los honorarios que perciban los abogados.[33] Cuando los postores presentan la misma calidad, el postor triunfante será la empresa que maneje el caso por la cuota más baja. Pero es improbable que los postores sean de la misma calidad. Una compañía que ofrece representar al grupo por una cuota de 1 millón de dólares en un caso en el que los intereses en juego suman 10 millones no le está

[33] Véase el análisis de estas subastas en In re Synthroid Marketing Litigation, 264 F.3d, 712 (7th Cir., 2001).

ofreciendo al grupo un mejor trato que otra compañía que ofrece representarla por 2 millones si la primera firma tiene sólo 50% de probabilidades de ganar mientras que la segunda tiene 75% de probabilidades de triunfar (caso en el que el beneficio esperado de representar a la clase es de 5 500 000 dólares, frente a sólo 4 millones de la primera firma). Aun así, habiendo ganado la puja, la compañía puede decidir trabajar menos de lo que el tribunal esperaba. Con unas probabilidades de 75% de tener la victoria, la compañía tiene una recompensa esperada de honorarios que suma 1.5 millones. Supongamos que el componente de ganancia es de 500 000 dólares y que, al dividir a la mitad su inversión de 1 millón tratando de ganar el caso, la compañía reducirá sus probabilidades de victoria de 75 a 60%. Su ganancia esperada será ahora más elevada: 700 000 dólares (60% de los 2 millones menos su nuevo costo de 500 000) *versus* 500 000.

El proceso de subasta puede ser estructurado para minimizar estos problemas. La puja se puede limitar a las compañías que tienen un historial, si bien esta política les pondrá las cosas difíciles a los que acaban de ingresar al mercado. O se pueden ofrecer bonos por resultados excepcionales. Esto provocará que las mejores firmas hagan una puja más baja, puesto que tendrán expectativas superiores sobre la obtención del bono.

Aun si no existen las subastas, la estructura de la compensación de honorarios en las demandas colectivas parece ser económicamente racional, aun cuando el nivel de los honorarios pueda ser demasiado alto. Las compensaciones son más elevadas cuando el riesgo de pérdida es más grande; tienden a ajustarse a la escala del resultado (con lo que se recompensa la producción acabada y no sólo la materia prima); y constituyen un porcentaje en declive de la cantidad de dinero recuperado para la clase.[34] Nótese que, entre más pronunciado sea el declive, mayor será la reivindicación del instrumento usado para la demanda colectiva; pues su empleo parte de la premisa, tal como vimos antes, de que existen costos fijos por la litigación —es debido a esos costos que el gasto realizado por litigar respecto a una reclamación pequeña es probablemente prohibitivo—, lo que implica que los costos promedio caen con el número y la homogeneidad de las reclamaciones que son agregadas en la demanda colectiva.

Hasta ahora he dado por sentado que, si bien las reclamaciones hechas por cada uno de los miembros del grupo son pequeñas, los daños agregados son de envergadura. Pero supongamos que son pequeños. Supongamos que el grupo está integrado por mil miembros, cada uno de los cuales tiene derecho a daños por 10 dólares. Por el bien de la disuasión, queremos que el demandado pague 10 000 dólares, pero ¿a quién? El costo de identificar a los miembros del grupo y verificar sus reclamaciones está llamado a superar los 10 000 dó-

[34] Véase Theodore Eisenberg y Geoffrey P. Miller, "Attorney Fees in Class Actions Settlements: An Empirical Study", 1 *J. Empirical Legal Stud.,* 27 (2004).

lares. En tales circunstancias, en ocasiones los tribunales otorgarán, en lugar del pago de daños, lo que llaman la ayuda *Cy pres*,[35] por analogía con la doctrina del *Cy pres* concerniente al derecho cuasidelictivo, y abordada en el capítulo XIX. Recuérdese que la doctrina permite que el dinero depositado en un fideicomiso sea derivado a un objeto diferente. La acción de clase análoga consiste en donar a una fundación de caridad el dinero que de otro modo les sería otorgado a los miembros del grupo como pago de daños. En nuestro ejemplo, el pago de 10 000 dólares a una organización de beneficencia cuya misión coincida, o al menos se traslape, con el interés del grupo, por ejemplo (si la demanda colectiva se basara en un estatuto de protección al consumidor) una fundación de caridad preocupada por proteger al consumidor, amplificaría el efecto del pago modesto de daños en protección de los consumidores al añadir al efecto disuasorio del fallo una inversión en otros medios usados para proteger a éstos contra el tipo de perjuicio causado por el demandado. Una fundación podría llevar a cabo alguna actividad para minimizar las violaciones al estatuto de protección al consumidor; los miembros de la clase, al recibir cada uno 10 dólares, no harían nada.

En el extremo opuesto deberíamos evaluar si alguna vez deberían permitirse las demandas colectivas cuando las reclamaciones individuales son grandes, como ocurre en las reclamaciones "masivas de cuasidelitos" que implican la muerte o lesiones graves, en vez de pequeñas. Por una parte, no existen los beneficios de la demanda colectiva que se tienen al permitir que una reclamación pequeña sea reivindicada, aunque todavía hay economías de escala y los costos de agencia son menores que en la demanda colectiva de reclamación pequeña porque cada miembro del grupo tiene un interés suficientemente grande para otorgarle un incentivo para el monitoreo del desempeño de los abogados en favor del grupo. Por otra parte, los costos del error son muy grandes porque son iguales a la probabilidad del error multiplicada por el costo del error si se comete. Un grupo integrado por las familias de personas muertas o gravemente lesionadas como una consecuencia presunta del fármaco defectuoso del demandado puede tener fácilmente reclamaciones de daños plausiblemente estimados en miles de millones de dólares, todas ellas dependientes del veredicto de un solo juez si se permite que el juicio proceda como una demanda colectiva. Aunque el juicio no tenga ningún mérito y la probabilidad de un veredicto erróneo en favor del grupo sea de sólo 2%, el demandante podría verse obligado a llegar a un arreglo para evitar el riesgo de la quiebra.

La práctica inglesa y de Europa continental de exigir a la parte perdedora de un juicio que rembolse (indemnice) los honorarios del abogado de la parte ganadora podría proveer aparentemente una alternativa para la demanda co-

[35] Véase la discusión en Hughes *vs.* Kore of Indiana Enterprise, Inc., 731 F.3d, 672 (7th Cir., 2013).

lectiva como un método de vindicación de las reclamaciones meritorias pequeñas. Por pequeña que sea la reclamación, el reclamante no se verá disuadido de emprender la persecución de sus remedios legales por el costo de la litigación puesto que sus gastos de litigación le serán rembolsados si gana. Pero hay varios problemas.

1. A menos que el demandante esté seguro de ganar, su costo de litigación esperado podría superar todavía el beneficio esperado para él. Si su reclamación es de 1 dólar, la probabilidad de que gane es de 90% y sus gastos de litigación son de 100 dólares, el beneficio esperado de la litigación será de sólo 90 centavos de dólar y el costo esperado de 20 dólares (suponiendo que los costos de litigación de su oponente sean también de 100, de modo que si el demandante pierde deberá un total de 200 dólares por concepto de honorarios del abogado). Por lo tanto, no demandará.

2. La indemnización carece de la característica de economías de escala de la demanda colectiva. Supongamos que hay 1 000 reclamaciones idénticas de 1 dólar cada una, el costo de litigación de cada una es de 100 y la probabilidad de ganar las reclamaciones es de 100%. Si los 1 000 reclamantes demandan —como podrían hacerlo, porque cada uno de ellos tiene una ganancia neta esperada de 1 dólar del juicio— se gastarán 100 000 dólares para vindicar tales reclamaciones. Si las reclamaciones se hubiesen agregado en una demanda colectiva, los gastos del juicio podrían haber sido sólo una pequeña fracción de esta cifra. El ejemplo es poco realista porque nunca hay 100% de certeza de ganar y porque la indemnización, siendo incompleta por las razones apenas mencionadas, no habrá de inducir a la litigación por intereses en juego tan pequeños. No obstante, esto solamente significa que en modo alguno se convocará a juicios de poca monta, lo que se traducirá en una pérdida de asignación (la conducta errónea no será objeto de disuasión). Existe aún una economía derivada del instrumento de demanda colectiva; eso es mejor que no tener ningún juicio y, asimismo, mejor que contar con muchas demandas pequeñas de carácter no económico.

Podría parecer que, si la demanda colectiva no fuese una alternativa viable, la indemnización conduciría a una cantidad de litigación socialmente excesiva porque se gastarían 100 000 dólares litigando reclamaciones que sólo valen 1 000. Sería muy probable, con todo, que la viabilidad de tales juicios hubiera disuadido al demandado de cometer el acto ilícito en primer lugar. Éste es un beneficio importante de la indemnización y también demuestra la falacia de hacer una comparación del costo de litigación de una reclamación con el valor preciso de esa reclamación.

La cuestión más debatida en el caso de una indemnización es su efecto sobre la tasa de litigación: quienes defienden la indemnización sostienen que es la respuesta para la crisis de la carga de trabajo de los juicios. Con ella, la condición para la litigación, la desigualdad (1) en XXII.5, se vuelve

$$P_p(J + C) - C - (1 - P_p)\, C + S > P_d(J + C) + C - (1 - P_d)C - S. \qquad (3)$$

La ganancia para el demandante, si gana el juicio (y la pérdida para el demandado si lo pierde) incluye ahora los gastos de litigación del demandante (*C*) y los daños otorgados (*J*); pero la ganancia esperada por el demandante debe reducirse por el gasto de litigación del demandado (también C) descontado por la probabilidad subjetiva del demandante de pérdida, 1 – Pp. Se requiere un ajuste similar para la pérdida esperada por el demandado de la litigación.

La desigualdad (3) puede rescribirse

$$(P_p - P_d)J > 2[(P_d + 1 - P_p)\, C - S]. \qquad (4)$$

La diferencia entre esta formulación y la desigualdad (2) (la condición paralela para la litigación sin indemnización) es el miembro derecho. Esta diferencia será mayor o menor en la desigualdad (4) que en la desigualdad (2) dependiendo de que $P_d + 1 - P_p$ sea mayor o menor a 1. Si es menor, el miembro derecho de (4) será menor que el miembro derecho de (2), lo que vuelve a la litigación más probable con un sistema de indemnización. Y será mayor a 1 sólo si P_d es mayor a P_p, es decir, sólo si el demandado estima la probabilidad del triunfo del demandante por encima de la estimación del propio demandante. Pero en estas circunstancias (pesimismo mutuo) se arreglará el caso de todos modos. Así que, para los fines de la evaluación del efecto de la indemnización sobre la probabilidad de litigación, podemos confinar nuestra atención a la clase de casos en los que P_p supera a P_d. En todos esos casos, la indemnización vuelve a la litigación *más* probable de lo que sería sin indemnización al volver al miembro derecho de (4) más pequeño que el de (2). La indemnización aumenta los intereses en juego (el demandante gana más si gana, el demandado pierde menos si *él* gana); y, como vimos antes, un incremento en los intereses que están en juego (*J*) aumenta las probabilidades de litigación.

3. La indemnización crea una litigación satelital, y ésta puede ser muy costosa. El costo mayor es el de determinar el costo razonable del esfuerzo de litigación del demandante. ¿Entonces por qué no darle el costo efectivo de su esfuerzo? El problema es que la indemnización internaliza en cierto sentido un costo externo, pero en otro sentido crea un costo. De acuerdo con la regla estadunidense, cada parte externaliza el costo de litigación de la otra parte. Según la regla inglesa y de Europa continental, cada parte externaliza su propio costo de litigación, porque si gana trasladará este costo a la otra parte. Por consiguiente, si espera ganar, la ganancia esperada neta que obtenga de ganar será más elevada si recupera igualmente sus costos de litigación. He ahí la razón por la que, si el "optimismo mutuo" es una condición para que un caso no llegue al arreglo, habrá menos arreglos bajo la regla inglesa y continental.

A fin de poner límites a esta externalidad, los tribunales otorgan sólo un honorario de abogado *razonable*. Los tribunales tratan de calcular cuánto habrían costado los servicios legales óptimos para un caso de esta clase en un mercado de servicios legales que opere sin externalidades, y eso se convierte en la cuota razonable otorgada por el tribunal. Se trata de un cálculo difícil y, por ello, incrementa el costo de la litigación satelital.

4. Es probable que la indemnización aumente la tasa de apelación —una consecuencia problemática porque, como veremos más adelante, los tribunales de apelación son el cuello de botella de un sistema judicial. La existencia de una indemnización otorga al apelante exitoso un beneficio adicional derivado del hecho de ganar su apelación—, al revertir el otorgamiento de honorarios de abogados a su oponente en el tribunal de instrucción. Sea p la probabilidad que tiene el apelante de ganar la apelación, J el fallo que obtendrá (o la terminación, si es el demandado y perdió en el tribunal de instrucción) si gana, C_l el gasto legal en que incurrió en la instrucción (y para simplificar suponemos que el gasto de su oponente fue el mismo) y C_a el gasto en que incurrirá en la apelación (que también suponemos igual al de su oponente). Si cada parte sufraga sus propios gastos, el beneficio esperado de la apelación para el apelante es simplemente $pJ - C_a$. Pero con una indemnización es $p(J + C_l + C_a) - C_a - (1 - p)\,C_a$. Esto es así porque, si el apelante gana (con probabilidad p), recupera no sólo el monto del fallo, sino también sus gastos legales en el tribunal de instrucción y el de apelación, mientras que si pierde (con probabilidad 1 – p) sólo se verá privado de sus gastos legales y los de su oponente en el tribunal de apelación. Supongamos, por ejemplo, que $C_l = 0.3J$, que C_a es sólo un quinto del tamaño de C_l (y, por lo tanto, igual a $0.06J$), y que p es 0.5. Entonces el beneficio esperado de una apelación para el apelante, si cada parte sufraga sus propios gastos legales, se simplifica a $0.5J - 0.06J = 0.44J$, mientras que el beneficio esperado bajo el traslado de los honorarios en ambos sentidos se simplifica a $0.5(J + 0.3J + 0.06J) - 0.06J - 0.5(0.06J) = 0.59J$ y es por lo tanto mayor.

Pero el análisis está incompleto.

1. Al aumentar la varianza de los posibles resultados de la litigación, la indemnización desalienta la litigación para quien siente aversión por el riesgo. Sin indemnización, el demandante recibe $J - C$ si gana el juicio y paga C si pierde; el intervalo de los resultados va así de $J - C$ a $-C$. Con indemnización, el intervalo es más amplio, de J a $-2C$. El análisis es similar para el demandado. ¿Pero cuál es exactamente la importancia de la aversión por el riesgo en la litigación? En la litigación comercial, es probable que la importancia no sea muy grande, debido a que las corporaciones pueden ser efectivamente neutrales respecto al riesgo si sus accionistas mantienen carteras diversificadas, tal como se explicó en el capítulo XV. Aun en la litigación por perjuicio personal, la aversión al riesgo puede carecer de importancia. El

riesgo del demandante puede protegerse mediante un contrato de honorarios contingentes y el demandado podría ser una corporación o un asegurado (o ambas cosas).

2. El costo para una de las partes de exagerar la probabilidad de su triunfo es mayor en un sistema de indemnización porque tiene más que perder si sale derrotado; un error es, por lo tanto, más costoso. Con el sistema estadunidense, su oponente toma para sí parte de los costos generados por la decisión de demandar o, sin ninguna base, defender un caso, a saber, los honorarios de su propio abogado.

Por lo tanto, la indemnización debe generar una convergencia mayor de las estimaciones de las partes acerca del resultado probable de la litigación, reduciendo así la tasa de litigación. Por cierto, esto sugiere una posible razón para que la indemnización sea la regla en Inglaterra pero no haya prendido en los Estados Unidos. La adhesión rígida al *stare decisis* por parte de los jueces ingleses, la mayor claridad de las leyes inglesas, que se debe al hecho de que el Parlamento es un organismo más disciplinado que la legislatura estadunidense (en efecto es unicameral, y está controlado por el poder ejecutivo: el gabinete), la virtual abolición del jurado civil en Inglaterra y la mayor sencillez del derecho inglés, atribuible a la ausencia de los estados y de una Constitución impugnable, vuelven los resultados de la litigación más previsibles en Inglaterra que en los Estados Unidos. Esto vuelve a su vez más culpable en Inglaterra un pronóstico errado acerca del resultado, en el sentido de que es más fácilmente evitable, con el resultado de que es más probable que la penalización de tales errores disminuya su incidencia allí que bajo un sistema como el nuestro, donde son inevitables muchos errores de pronóstico debido a la incertidumbre de la litigación. Los códigos estatutarios detallados, los jueces profesionales y la ausencia de jurados que caracterizan a los sistemas jurídicos continentales podrían explicar la razón de que la indemnización sea la regla en la Europa continental al igual que en Inglaterra.

3. Incluso bajo la ley estadunidense, una parte que gana puede obtener la indemnización de sus honorarios si la reclamación (o la defensa) de la otra parte resulta ser totalmente carente de bases (frívola). Podría parecer que no se necesitaría ningún castigo aun en el caso frívolo, ya que el costo de rechazar una oferta de arreglo y luego perder en el juicio es soportado por la persona que comete el error. Pero no es así del todo. La otra parte soporta también los costos de la litigación, y, como sabemos, también el contribuyente soporta algunos de ellos. Estos costos son externos al perdedor.

Al incrementar los intereses en juego tal como los perciben las partes, la indemnización debería incrementar la tasa de litigación; y existen evidencias de que así ocurre.[36] También debería acrecentar la precisión con que procede

[36] Edward A. Snyder y James W. Hughes, "The English Rule for Allocating Legal Costs: Evidence Confronts Theory", 6 *J. Law, Econ. & Organization,* 345 (1990).

la litigación al incrementar la cantidad gastada por cada caso; alienta al litigante optimista a gastar mucho en litigación porque espera que sus costos sean sufragados en última instancia por su oponente, aunque esta tendencia podría frenarse por la negativa judicial a otorgar más que un honorario razonable. Y, en la medida en que la litigación deriva de estimaciones descuidadas sobre los resultados, la indemnización, al penalizar esos descuidos, puede impedir el desarrollo de una serie de litigios en suficiente grado para reducir la tasa de litigación hasta alcanzar un equilibrio.

La indemnización alienta la presentación de reclamaciones pequeñas. El efecto sobre el número de reclamaciones meritorias podría ser dramático si, por ejemplo, la regla de responsabilidad fuese la responsabilidad estricta, no les conviniera a los demandados potenciales reducir su exposición a la responsabilidad reduciendo sus niveles de actividad y muchas de las reclamaciones potenciales fuesen por cantidades de dinero muy pequeñas. Habría más reclamaciones porque, podríamos darlo por sentado, los demandados potenciales no se verían disuadidos de observar la conducta que originó tales reclamaciones por la perspectiva de tener que pagar. ¿Pero *muchas* más reclamaciones? Eso es dudoso, porque la perspectiva de recuperar una pequeña cantidad (una vez restados los honorarios de los abogados) en una demanda no constituye una motivación fuerte. De modo que un mayor efecto que habría de tener sería el de desalentar las demandas por perjuicios fundadas en reclamaciones frívolas. Sin la indemnización, el costo neto esperado de la litigación de una reclamación que ambas partes creen frívola es la misma para el demandante y el demandado; y, en virtud de que *hay* un costo para el demandado, el demandante, por analogía con el juego del gallina, podría ser capaz de obtener algo del demandado por vía del arreglo. Con indemnización, el costo neto esperado de la litigación de tal reclamación se duplica para el demandante pero baja a cero (o al menos a cerca de cero) para el demandado. El demandante sabe que perderá (si el caso se lleva a juicio) y tendrá que cubrir los gastos legales del demandado (C), además de correr con los propios (también C, para un total de $2C$). El demandado sabe que ganará y que sus gastos legales serán rembolsados, de manera que la defensa del caso no le costará un centavo.

Pero no todos los juicios de perjuicios son frívolos. Un juicio es frívolo si no tiene ninguna base en la ley, lo que implica una probabilidad muy baja de que el demandante gane si la demanda se litiga. Un juicio de perjuicios podría ser un juicio meritorio en el que los daños potenciales fuesen tan ligeros que el valor esperado neto del juicio (es decir, restando los costos de litigación del demandante), si se litigara, sería negativo. ¿Deberíamos preocuparnos por los juicios de perjuicios, frívolos o no, o es éste otro ejemplo en que la teoría de juegos sólo nos dice que *podría* haber tal cosa? La naturaleza secuencial de los gastos de litigación, que como vimos antes podría conducir a gastos mayores de lo que está en juego en el caso, sugiere que hay en efecto un problema

de juicios de perjuicios.[37] Supongamos que el fallo neto esperado para el demandante es de 2 000 dólares, y que si litiga el caso le costará 3 000. Supongamos que la litigación tiene dos etapas, en cada una de las cuales tendrá que gastar el demandante 1 500 a fin de llevar el caso adelante. Si gasta los primeros 1 500, este gasto se convierte en un costo hundido, irrecuperable, que no influirá sobre sus decisiones de gasto subsecuentes si el demandante es racional. Por lo tanto, gastará los siguientes 1 500, porque al hacerlo así generará una ganancia neta mayor (2 000, el valor esperado del juicio si gasta un total de 3 000 dólares). Sabiendo esto, el demandado transigirá en la segunda etapa. Sabiendo que el demandado transigirá en la segunda etapa, es racional que el demandante gaste 1 500 en la primera etapa. (Éste es otro ejemplo de una estrategia racional de compromiso previo.) Sabiendo *esto*, el demandado transigirá en la primera etapa, donde el intervalo de negociación (antes de que el demandante gaste los primeros 1 500) es de 0 a 2 000 dólares (omitiendo los gastos de litigación del demandado, que aumentarán el intervalo). Pero el acusado puede tener la capacidad para ganarle este juego al demandante realizando su propio gasto en el mismo momento, lo que reducirá su costo marginal de la defensa.

Lejos de eliminar la demanda por perjuicio fundada en un reclamo meritorio, la indemnización la convierte en un juicio del valor esperado neto. Supongamos que en nuestro ejemplo anterior se generó el fallo esperado de 2 000 dólares por una probabilidad de 100% de que, si el caso se fuera a juicio, el demandante recibiría un fallo de 2 000. Entonces, el siguiente valor neto esperado del juicio para él sería positivo; en efecto sería de 2 000 dólares, puesto que el demandado tendría que pagar al demandante, también con una probabilidad de 100%, 3 000 de costos de litigación. En otras palabras, el fenómeno de la demanda por perjuicio fundamentada en un reclamo meritorio es un artefacto de la regla estadunidense, el cual niega la indemnización.

La relación entre la litigación depredadora y las demandas por perjuicios es, así, que aquélla está diseñada para alterar el comportamiento del demandado, en tanto que la litigación por perjuicios está diseñada para lograr un arreglo, aunque el valor esperado neto para el demandante por litigar el caso en un juicio es negativo.

XXII.13. La Regla 68 y la indemnización unilateral

La Regla 68[38] de las Reglas Federales del Procedimiento Civil dispone que, si el demandado hace una oferta de arreglo antes del juicio, y si el demandante

[37] Véase Lucian Arye Bebchuk, "A New Theory Concerning the Credibility and Success of Threats to Sue", 25 *J. Legal Stud.,* 1 (1996).

[38] Analizada desde un punto de vista económico en George L. Priest, "Regulating the Content and Volume of Litigation: An Economic Analysis", 1 *Supreme Ct. Econ. Rev.,* 163 (1982); Geoffrey

la rechaza y luego le va peor en el juicio, el demandante deberá pagar sus propios costos del juicio, aunque como parte ganadora tendría normalmente derecho a insistir en que el demandado pague los costos de ambas partes.[39] ¿Aumenta la Regla 68 la tasa de arreglos? Es probable que no. Esa regla vuelve la litigación más costosa para el demandante porque, aun si gana, podría tener que pagar sus propios costos (incluyendo a veces los honorarios de su abogado) y, por ende, demandará menos en el arreglo del caso. Pero por la misma razón el demandado, quien ahora tiene menos que perder de la litigación, ofrecerá menos. Por lo tanto, el intervalo del arreglo no será mayor. Pero la regla tiene un efecto indirecto que podría alentar el arreglo. Aumenta el incentivo del demandado para hacer un ofrecimiento realista, es decir, cercano a la ganancia que el demandante espera obtener del juicio, porque si el ofrecimiento no es realista y se rechaza, y el caso se va a juicio y el demandante gana, el demandado se beneficiará del hecho de tener una oferta de arreglo sólo si al demandante le va peor en el juicio en relación con lo que obtendría si hubiera aceptado el ofrecimiento.

En algunos casos le resultará imposible al demandado formular un ofrecimiento de la Regla 68 que supere el fallo del juicio si el caso se va a juicio y el demandante gana. Supongamos, por ejemplo, que el demandante cree que tiene una probabilidad de 50% de perder y una probabilidad de 50% de ganar 100 000 dólares y, por lo tanto, no aceptará un ofrecimiento de arreglo menor de 50 000 (podemos omitir el gasto de la litigación y el arreglo en este ejemplo). El demandado cree que el demandante tiene una probabilidad de triunfo de sólo 40%, de modo que no ofrecerá más de 40 000. Éste podría ser un buen ofrecimiento porque la estimación por parte del demandado de las probabilidades del demandante es mejor que la estimación del demandante, pero no hay modo de que el demandado pueda obtener del ofrecimiento un beneficio de la Regla 68. Si el demandante pierde, la Regla 68 no opera (porque entonces el demandante tiene que pagar sus costos de todos modos, como la parte perdedora); si gana, ganará 100 000 dólares, una suma mayor que la del ofrecimiento del demandado.

El efecto principal de la Regla 68 es el de transferir riqueza de los demandantes a los demandados, pero podría no ser mucho. El demandante no rechazaría un ofrecimiento de la Regla 68 a menos que esperara que le fuera

P. Miller, "An Economic Analysis of Rule 68", 15 *J. Legal Stud.,* 93 (1986); David A. Anderson, "Improving Settlement Devices: Rule 68 and Beyond", 23 *J. Legal Stud.,* 225 (1994); Tai-Yeong Chung, "Settlement of Litigation under Rule 68: An Economic Analysis", 25 *J. Legal Stud.,* 261 (1996).

[39] En ausencia de una ley, los "costos" a los que se refiere la regla son los rubros más bien menores que aun bajo la regla estadunidense debe rembolsar de ordinario el perdedor al ganador, incluidos los honorarios del tribunal, los costos de las copias y (algunos) honorarios de los testigos, pero excluyendo el rubro grande: los honorarios de los abogados. Se ha sostenido que la regla cubre los honorarios del abogado cuando la ley establece tales honorarios, por lo menos si la ley llama costos a tales honorarios.

mejor en el juicio; y, si le fuera peor que lo esperado, ello significaría que había cometido un error al rechazar el ofrecimiento. Por lo tanto, la regla penaliza al demandante sólo por un error. Pero esto omite la aversión al riesgo. La decisión de aceptar o rechazar el ofrecimiento de arreglo tiene un riesgo mayor en sentido descendente cuando está operando la Regla 68, sobre todo cuando se aplica la regla a los honorarios del abogado tanto como a los costos convencionales. Por otra parte, la Regla 68 aumenta el incentivo del abogado para hacer un ofrecimiento de arreglo realista, y el ofrecimiento será más generoso entre mayores sean los beneficios conferidos por la Regla 68 al demandado si el demandante rechaza el ofrecimiento y luego le va peor en el juicio. Por lo tanto, es posible que la regla sea sólo ligeramente desventajosa para los demandantes. Con todo, tendría sentido enmendar la regla a fin de permitirle a un demandante que hiciese una oferta de arreglo que tuviera el mismo efecto que la oferta del demandado.

La Regla 68 favorece al demandado; las leyes (que son comunes) que otorgan sólo a un demandante triunfante un derecho al rembolso de sus honorarios legales favorecen al demandante. Cambian la desigualdad (3) a

$$P_p(J + C) - C + S > P_d(J + C) + C - S, \tag{5}$$

que puede rescribirse como

$$(P_p - P_d)J > 2[(P_d/2 + 1 - P_p/2)\, C - S]. \tag{6}$$

En todos los casos en los que $P_p > P_d$, la situación de optimismo mutuo que de ordinario es necesaria para que ocurra la litigación, una comparación con la desigualdad (4), la condición paralela de la regla inglesa y de Europa continental, revela que la litigación es menos probable. Esto es así porque la cifra negativa en el miembro derecho, $P_p - P_d$, se reduce a la mitad. Pero una comparación con la desigualdad (2) revela que la litigación es más probable que con la regla estadunidense. Los intereses en juego para el demandante son los mismos que con el traslado de los honorarios, pero son menores para el demandado: tiene menos que ganar de la litigación que con el traslado de los honorarios y, por lo tanto, tendrá menos incentivos para litigar. Esta conclusión es contraria a la opinión de que una indemnización unilateral fomenta la litigación más que otra regla. Además, podría no generar más litigación que la regla estadunidense (de no indemnización). Un estudio descubrió que la promulgación en 1976 de la ley de traslado unilateral de los honorarios en los casos de derechos civiles no había aumentado el número de tales casos.[40] La ley aumentó el incentivo de los demandados potenciales para arreglar las

[40] Stewart J. Schwab y Theodore Eisenberg, "Explaining Constitutional Tort Litigation: The Influence of the Attorney Fees Statute and the Government as Defendant", 73 *Cornell L. Rev.*, 719, 756-758 (1988).

reclamaciones de derechos civiles antes de que se iniciara la litigación y para evitar las violaciones efectivas o incluso presuntas (o por lo menos detectables) de los derechos civiles de demandantes potenciales. Presumiblemente, estos efectos contrarrestan el efecto de la ley al reducir la tasa de arreglos; si se establecen menos juicios, una tasa de arreglos más baja (en una instancia de instrucción superior) es congruente con una cantidad más baja de juicios de instrucción.

XXII.14. La *RES JUDICATA* y el impedimento colateral

Podría parecer sorprendente la negativa de los tribunales a permitir que la misma reclamación vuelva a litigarse entre las mismas partes *(res judicata)*. Habiendo perdido una vez, es de suponerse que la parte se resistiría a intentarlo de nuevo; la pérdida es una prueba del resultado probable de su segundo juicio. ¿Pero por qué debiera prohibírsele el intento, así como a un anunciante no se le prohíbe que repita una campaña publicitaria que fracasó la primera vez apenas unos meses atrás? La respuesta es que el costo de la nueva litigación es positivo, mientras que el beneficio de la reducción de los costos de los errores es, en general, cero puesto que no hay manera de determinar cuáles resultados de una serie de resultados incongruentes (A demanda a B y pierde; A demanda a B de nuevo y gana; B demanda ahora a A para recuperar el fallo de A contra B y gana, etc.) son los correctos. Pero esto es a condición de que los intereses en juego sean los mismos o similares en las diversas acciones y (algo relacionado) el primer tribunal se encuentre aproximadamente en el mismo plano de competencia que los tribunales subsecuentes (por ejemplo, no un tribunal de tráfico). Ambas condiciones son necesarias para volver improbable que los costos del error en una segunda litigación fuesen menores que los de la primera. La primera condición puede darse por sentada; la *res judicata* inducirá a ambas partes a invertir adecuadamente en la primera demanda, a sabiendas de que el fallo en ella le cerrará el paso a una litigación ulterior. La segunda condición ha dado nacimiento a la regla de que la *res judicata* no se aplicará si la primera demanda se presentó ante un tribunal o una agencia administrativa que carece de competencia institucional para hacer determinaciones definitivas.

La *res judicata* (cosa juzgada) prohíbe también que un demandante "fragmente" su reclamación. Supongamos que un demandante tiene una causa de acción de cuasidelito y una de contrato contra el demandado derivadas del mismo incidente. No se le permitirá que demande primero por una causa de acción y luego por la otra. Aunque hay diferentes causas de acción, se considerarán como una sola reclamación en reconocimiento de las economías obtenibles mediante la combinación de las dos teorías de responsabilidad en un solo juicio.

La *res judicata* puede ser algo menos importante en un sistema como el inglés donde las reclamaciones por perjuicio se penalizan mediante un traslado de honorarios a cabalidad. Si todos los costos de litigación se internalizaran (incluidos todos los costos del sistema judicial y la totalidad de los costos de cola de otros litigantes), tal vez podríamos vivir sin una regla de *res judicata*; quizás un juicio de repetición se presentaría solamente si el fallo de la primera demanda hubiese sido erróneo. No obstante, dado un alto nivel de error judicial que resulta inevitable y, por consiguiente, el papel inevitable del azar en los fallos, la *res judicata* es necesaria para acabar con la litigación redundante que no tiene ningún valor social neto esperado.

Una interrogante interesante es ésta: ¿cuándo podrá usarse un fallo para prohibir la nueva litigación de las mismas cuestiones en una litigación subsecuente (el impedimento colateral), no necesariamente con la misma parte? Supongamos que A demanda a B y gana, y luego demanda a C, y cierta cuestión decidida en favor de A en su juicio contra B (quizá si un producto vendido por B y C está defectuosamente diseñado) se plantea también en el juicio de A contra C. Si se permitiera a A el uso del fallo anterior para impedir la nueva litigación de esta cuestión en su demanda contra C, se podría llevar a A a invertir recursos excesivos en el triunfo sobre esa cuestión en su juicio contra B. Por ejemplo, podría seleccionar como su primer demandado (esto es, como B) a alguien cuyo interés en la determinación correcta de la cuestión fuese demasiado pequeño para justificar la inversión de recursos considerables para lograr que se decida en su favor, mientras que A gastaría mucho, anticipando beneficios en litigaciones subsecuentes. Empero, sería relevante si pudieran intervenir C y otros demandados potenciales en la demanda de A contra B; pues, en principio, la intervención proporcionaría un método para consolidar todas las litigaciones potenciales a través de una.

Invirtamos ahora los hechos: B gana en el juicio de A contra él y trata de usar el fallo para cancelar una cuestión decisiva en un juicio similar de E en su contra. El problema discutido antes es un poco menos agudo porque es de presumirse que B no decidió quién lo demandaría primero. Pero todavía hay un peligro de que B pueda gastar recursos desproporcionados en el juicio de una cuestión, anticipando beneficios del hecho de poder usar el fallo en litigaciones subsecuentes.

Casos como éstos han causado pequeñas dificultades a los tribunales.[41] Pero supongamos ahora que A demanda a B y gana, tras de lo cual E promueve un juicio similar contra B y trata de usar el fallo de A para impedir que B vuelva a litigar cualesquiera cuestiones comunes.[42] Esto se llama im-

[41] Aun cuando sea por una razón expresa que resulta diferente de la sugerida en el texto: que la parte a la que se buscó para tener un impedimento colateral de un problema (C en el primer ejemplo, E en el segundo) nunca dejó de correr con gran suerte en el tribunal.

[42] De manera que B ha dejado de correr con suerte en el tribunal, en lo concerniente a la demanda que le entabló A.

pedimento colateral ofensivo. A menos que haya colusión entre A y E, no tendremos que temer que A invierta excesivamente en la prosecución de su juicio contra B porque espere utilizar ese fallo contra otros demandados; no lo hará. El peligro es más bien que B invierta excesivamente en su defensa contra la demanda de A, porque la consecuencia de la pérdida de ese juicio podría ser una responsabilidad mucho mayor en juicios subsecuentes. Por supuesto si B, habiendo invertido excesivamente en el primer juicio, lo pierde, ésa es una buena prueba de que la reclamación de A era buena: ¿y entonces por qué debe E probar de nuevo sus elementos esenciales? Pero si se permite que E utilice el fallo de A en esta forma, B tendrá, como hemos dicho, un gran incentivo para invertir desproporcionadamente en la defensa del juicio de A, y esto aumentará el peligro de un fallo erróneo en favor de B en ese juicio.

Consideremos ahora el impedimento colateral defensivo. A, quien tiene reclamaciones similares contra F, G y H, demanda primero a G y pierde. ¿Debería permitirse que F y H utilizaran el fallo de G para impedir las reclamaciones de A contra ellos? Presumiblemente, A escogió su caso más fuerte para plantearlo primero; ya que, aun cuando una victoria en el primer caso no habrá de tener un impedimento colateral para A, alentará a otros demandados a obtener un arreglo dando al menos cierta información, favorable a las perspectivas de A, en lo concerniente al resultado probable de sus juicios; y puede sentar un precedente útil para A en virtud de que cualesquiera prescripciones legales que haga el juez en su caso pueden ser citadas y seguidas por los jueces en casos subsecuentes. Si pierde ese primer caso, esto implica que los casos restantes carecen también de mérito. Pero ahora surge la imagen refleja del problema que acabamos de discutir en conexión con el impedimento colateral ofensivo. A podría utilizar muchos recursos en ese primer juicio, sabiendo que una pérdida allí será un desastre. Lo que tiene en juego B es algo mucho menor. Esta asimetría podría permitir que A ganara un caso que no merece ganar. La probabilidad de tal resultado aumenta si un fallo en favor de B, cuando lo hay, puede utilizarse contra A por demandados subsecuentes.[43]

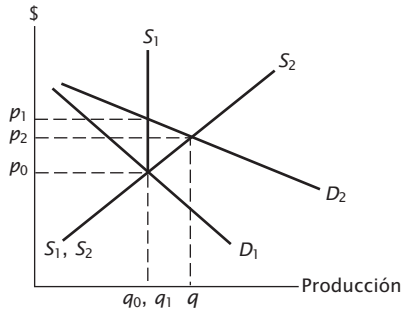
[43] Este análisis sugiere una objeción a la *stare decisis* en algunas clases de casos, si bien se trata de una objeción menos contundente. Un caso en el que las partes han realizado una inversión relativamente escasa puede generar un precedente que habrá de controlar el desenlace de casos subsecuentes en los que las partes podrían haber invertido muchos más esfuerzos en prestar ayuda al juez con base en datos apegados a fin de dar lugar a una decisión sólida. Este efecto se mitiga con la práctica de tratar a la *stare decisis* como una presunción en vez de como un requisito absoluto. En un caso posterior, el juez puede relegar el precedente que estuviera basado en una inversión inadecuada realizada por las partes en el caso anterior y que, por lo tanto, resultara presumiblemente de calidad inferior.

XXII.15. La demora judicial y las crisis de la carga de trabajo

La demora de la ley ha sido motivo de lamento en la literatura popular desde la época de Shakespeare, pero gran parte de la crítica tradicional sobre esa demora es superficial. La relación inversa entre el costo y el tiempo (véase antes X.8) implica que la eliminación del intervalo completo que media entre el inicio y la decisión de un juicio sería ineficiente. Además, la demora judicial es una cola "figurativa" por oposición a una "literal". Cuando se hace una fila de espera por una mesa en un restaurante se forma una línea "literal"; ésta impone un costo de oportunidad medido por el valor del tiempo del cliente mientras espera. No hay tal costo en la demora judicial, ya que (si omitimos la detención previa en los casos penales) los litigantes están en libertad para ocuparse de sus asuntos mientras esperan el juicio. Con todo, hay otros costos de la demora judicial indebida; por ejemplo, la degradación de las pruebas, que incrementa los costos de errores legales; y la mayor incertidumbre generada cuando el juicio se pospone por largo tiempo, lo que aumenta los costos (a la vez que reduce la probabilidad) de un arreglo.

La demora judicial excesiva no es, como se sugiere a veces, la consecuencia inevitable del hecho de que la demanda de litigación sea grande y la cantidad de tiempo de los jueces sea limitada. La demanda de langostas es grande también, y la capacidad para expandir la producción a fin de satisfacer nuevos aumentos de la demanda es también limitada. La gente hace fila para comprar litigación, pero no para comprar langostas, porque el tiempo judicial no se raciona por el precio, lo que sí ocurre con las langostas. Si la demanda de langostas aumentara más deprisa que la oferta, el precio aumentaría hasta que se igualaran la demanda y la oferta. Un sistema apropiadamente graduado de recargos para quienes deseen que sus casos se escuchen pronto tendría el mismo efecto sobre la litigación. Si fuesen muy elevados los precios necesarios para vaciar el mercado (eliminar la fila), ello sería una señal de que probablemente se justificaría por su costo una inversión de recursos en la contratación de más jueces. Los precios podrían no ser elevados. Sólo una fracción pequeña de los litigantes podría tener interés suficiente en un juicio temprano para pagar un recargo. Ésa sería una señal para no contratar más jueces.

La demanda de los servicios judiciales, tanto estatales como federales, ha venido creciendo rápidamente desde alrededor de 1960 (de manera vertiginosa en el periodo que va de la década de 1960 a la de 1980, con más lentitud desde entonces), pero no se ha hecho ningún esfuerzo para moderar la demanda y guiar la respuesta de la oferta. La respuesta principal al crecimiento de la demanda ha sido la contratación de más jueces y el apoyo al personal judicial. Es improbable que tal respuesta tenga un efecto considerable sobre la demora judicial, excepto a muy corto plazo. Al aumentar la calidad del remedio legal, por lo menos para quienes aprecian la justicia rápida, una ex-



GRÁFICA XXII.3

pansión del número de jueces inducirá a utilizar los tribunales a algunos individuos que antes habían sido disuadidos por la demora. La analogía es con la construcción de una carretera nueva para aliviar el congestionamiento del tráfico. La nueva carretera podría inducir a algunas personas que antes utilizaban otros métodos de transportación en vista del costo del congestionamiento a utilizar ahora el automóvil, hasta que la carretera esté casi tan congestionada como los caminos que sustituyó.[44] En ambos ejemplos, al aumentar la oferta en una forma que reduce el precio ajustado por la calidad, el gobierno aumenta simultáneamente la cantidad demandada.

La gráfica XXII.3 representa los efectos de un aumento inesperado de la demanda en un mercado privado. A corto plazo, cuando la oferta está fija, el aumento de la demanda (de D_1 a D_2) causa un gran aumento del precio, de p_0 a p_1. Pero a largo plazo, cuando los productores pueden expandir la capacidad para satisfacer la nueva demanda (razón por la cual la curva de oferta a largo plazo, S_2, se encuentra por debajo de la curva de oferta a corto plazo, S_1), el precio baja de p_1 a p_2. No baja de nuevo hasta p_0 porque los esfuerzos de los productores para atraer de otras industrias mediante una puja los insumos que necesitan hacen que los precios de tales insumos aumenten. En otras palabras, la elasticidad de la oferta a largo plazo no es infinita (una línea horizontal a la derecha de la intersección entre p_0 y q_0), porque algunos de los insumos utilizados en la elaboración del producto en cuestión son inherentemente escasos en relación con su demanda.

Estos problemas han moldeado la forma que ha adoptado la expansión del poder judicial para satisfacer la creciente demanda. El número de jueces ha aumentado, pero bastante menos que el número de casos. Los miembros del personal de apoyo (asistentes jurídicos, internos y externos, jueces magistrados, sección de cibernética) se han agregado a un ritmo mucho más veloz que

[44] "Casi" porque la oferta es ahora más nutrida, de modo que, si una carretera llegara a estar tan congestionada como el camino más pequeño al que remplazó, esto implicaría que había más personas que antes dispuestas a tolerar ese grado de demora… y ¿por qué habría de ser que la construcción de la carretera hubiese alterado los gustos de esa manera?

los jueces. Ésta puede ser la razón de la pronunciada caída en el número de casos civiles a los que ha puesto fin el tribunal de instrucción frente a otros métodos para darles cierre, como los arreglos y las concesiones de juicios sumarios. La elasticidad cruzada de la oferta de jueces y el personal de apoyo es mucho más grande en los cierres realizados fuera de los tribunales de instrucción que dentro de éstos, porque un juez no puede delegar su papel de presidir un juicio, pero sí puede delegar la redacción de las órdenes de un juicio sumario y la función de encabezar la mediación del arreglo al personal de apoyo. De este modo, la expansión en el número de casos le permite a un juez dar cierre a una cantidad mayor de éstos sin juicio de por medio y, así, hace posible que un volumen cada vez mayor de casos se procesen sin que haga falta un incremento proporcional del número de jueces.

Una peculiaridad del personal judicial consiste en que los jueces enfrentados a la serie de casos que constituyen la carga más pesada, a saber, los jueces de instrucción (aquellos que presiden sobre el proceso de litigación desde la presentación de una solicitud hasta el juicio, a diferencia de los jueces que presiden sobre las apelaciones hechas al fallo en el tribunal de instrucción) tienen la razón más baja de personal a juez, en tanto que los jueces con la carga más ligera de trabajo tienen la razón más elevada. Los juicios distritales federales lidian con la carga más pesada de casos en el sistema judicial federal, pero cuentan para su auxilio con el personal más esquelético; los magistrados de la Corte Suprema manejan la carga más ligera de casos pero se les depara el personal más numeroso (ponderado por la calidad: los jueces supremos tienen por lo general el mismo número de asistentes jurídicos, cuatro, que los jueces del tribunal federal de apelaciones, quienes integran el tercer nivel ubicado entre los jueces distritales y los jueces de la Corte Suprema, pero los asistentes jurídicos de esta última son por lo común más experimentados y han sido sometidos a un proceso más riguroso de selección que los asistentes del tribunal de apelaciones). En otros términos, en el sistema judicial federal se suele asignar personal con base en el rango del juez que lo habrá de emplear, no en la necesidad. Por supuesto, eso revela ineficiencia. Asimismo, entre más nutrido y más capacitado esté el personal, será menor la atención que deban prestar las autoridades responsables de la designación a la calidad de los jueces por designar, y, de esta manera, es de esperar que la política desempeñe un papel más destacado a la hora de hacer nombramientos.

En la búsqueda de una mejor forma de rebajar la demanda de tiempo de los tribunales que la de incrementar el número de jueces o el tamaño y la calidad de su personal, ¿habría uno de acogerse a la medida de imponer requisitos más estrictos en materia de la cantidad mínima en la controversia[45] o a la de fijar mayores cuotas por derechos sobre la presentación? El economista

[45] En la actualidad, un demandante no puede promover un juicio de diversidad (véase después XXVII.3) en un tribunal federal a menos que lo que esté en juego exceda de 75 000 dólares.

prefiere la última. El enfoque de la cantidad mínima en la controversia equivale a un derecho de presentación infinito para los casos por debajo del mínimo y un derecho de presentación de cero para los casos por encima del mínimo, mientras que un derecho de presentación fijo operaría como un impuesto proporcionalmente declinante sobre los juicios legales. Un derecho de 1 000 dólares, por ejemplo, sería un impuesto de 100% a un caso en el que estuviesen en juego 1 000 dólares, y un impuesto de 1% a un caso de 100 000. Si el derecho se fijara igual al costo del juicio para el sistema judicial (incluidos no sólo los costos directos sino también los costos de causar el apoyo de otros casos), confrontaría al litigante (presumiblemente el demandante, pero podría exigirse al demandado que rembolsara al demandante si éste ganara) con el costo social completo del uso del sistema. Un requerimiento de cantidad mínima en controversia no hace esto.

Ambos requerimientos pueden ser criticados por favorecer a los ricos. Pero la crítica es incompleta, aun si lo que está en juego y la riqueza tienen una fuerte correlación positiva. Dado que las partes gastarán más en litigación entre mayor sea lo que esté en juego en el caso, y los mayores gastos en litigación reducirán los costos del error, toda la sociedad, ricos y pobres por igual, tiene (dando por descontado que la aplicación rigurosa de la ley se realiza en el interés público… lo cual puede no acontecer si el cuerpo estatutario está conformado en su mayor parte por malas leyes) un interés en que se asignen los recursos judiciales de la más alta calidad a los casos mayores, dejando la instrucción de los casos menores en manos de los tribunales de menor calidad. Y, aun cuando los derechos de presentación pudieran parecer inservibles cuando los litigantes padecen indigencia, esto es incorrecto. La indigencia es meramente un argumento a favor de la discriminación de los precios, de la práctica de solicitarle al indigente cualquier pago de derechos que esté a su alcance. Esto es lo que se lleva a cabo en la litigación de los prisioneros, forzándolos a realizar un trueque entre los beneficios derivados de la litigación y los costos implicados en la reducción de dinero extra a gastar en la comisaría.

Un sistema de derechos de presentación plenamente compensatorios eliminaría el subsidio para la litigación que resulta de no cobrar a los litigantes los costos de operación del sistema judicial mismo, un subsidio que podría justificarse por los beneficios externos que confiere la litigación al crear reglas de conducta para la sociedad. Los requerimientos de la cantidad mínima en la controversia preservan el subsidio, porque si un caso satisface el requerimiento los litigantes no tienen que contribuir a los costos del sistema judicial. Pero el enfoque del subsidio podría preservarse en un sistema de derechos de presentación simplemente restando el subsidio óptimo de un derecho calculado para cubrir el costo pleno del juicio para el sistema judicial.

XXII.16. JURADOS Y ÁRBITROS

Las colas en los tribunales son casi siempre mayores para las partes que buscan juicios civiles por jurados. Tales juicios son más costosos que los juicios sin jurado a causa de sus honorarios (que, como veremos más adelante, subestiman sus costos sociales verdaderos) y porque un jurado se tarda normalmente más que un juez para instruir un caso a causa del tiempo invertido en seleccionar a sus miembros, además del tiempo más largo que toma explicarles los hechos a quienes no son expertos, aparte del tiempo invertido en la comunicación que se establece entre el juez y los integrantes del jurado. Por lo tanto, se "cobra" a las partes más por los juicios de este tipo al hacérseles esperar más tiempo en la fila.

Los Estados Unidos son ahora el único país grande donde se utiliza regularmente un jurado en casos civiles, y pudiera ser dudoso que los costos adicionales del juicio se contrarresten por una disminución, si hay alguna, de los errores en la determinación de los hechos (o en el número de sobornos, puesto que es más difícil sobornar a un jurado que a un juez, aunque, por lo demás, el soborno de jueces no es un problema serio en el sistema legal moderno). Pero el jurado ofrece un campo rico para el análisis económico.[46]

Una vez que se tomó la decisión de encomendar la determinación de los hechos a los legos, resultaba inevitable que se emplearan varios legos, en lugar de uno solo, para este propósito. Ciertamente esto incrementa los costos del jurado al aumentar el número de personas que tienen que ser remuneradas, la cantidad de tiempo que gastarán deliberando y la probabilidad de un jurado empatado, así como, por lo tanto, de un nuevo juicio de instrucción. Pero tener personas de variados antecedentes y puntos de vista deliberando juntas es algo necesario para contrarrestar la mayor pericia del juez profesional, para quien el jurado es un sustituto en la aclaración de los hechos y la aplicación de la ley a ellos. Esto sugiere el intercambio esencial de ventajas y desventajas en el diseño de un sistema de jurados económicamente óptimo. El agrandamiento del jurado reduce los costos del error al aumentar la diversidad de experiencia y capacidad que estará implicada en el proceso de determinación de los hechos: por ejemplo, al volver más probable que el jurado incluya por lo menos un miembro perspicaz y elocuente que pueda concentrar sus deliberaciones.

Y, como cualquier forma de la diversificación, incrementar el tamaño del jurado reduce el riesgo de los resultados extremos (siempre que la selección

[46] Alvin K. Klevorick y Michael Rothschild, "A Model of the Jury Decision Process", 8 *J. Legal Stud.*, 141 (1979); Klevorick, "Jury Size and Composition: An Economic Approach", en *The Economics of Public Services* (Martin S. Feldstein y Robert P. Inman [comps.], 1977); Donald L. Martin, "The Economics of Jury Conscription", 80 *J. Pol. Econ.*, 680 (1972).

sea aleatoria o poco menos). Esto tiene la implicación, quizá contraria a la lógica elemental, de que un demandado debería preferir a un jurado de 12 personas que a uno de 6. Entre más miembros de un jurado haya, más aleatoria será probablemente su composición y, por consiguiente, será menor la probabilidad de dar fallos extremos. En virtud de que un veredicto debajo de cero no es posible, todo lo que aminore el intervalo de los veredictos posibles favorece a los acusados.

Otra variable en la designación del jurado es la mayoría requerida para un veredicto. Una regla de unanimidad será más costosa que una de mayoría simple. Se requiere más tiempo para negociar un veredicto unánime —se necesita el acuerdo de más partes—, y así aumenta la probabilidad de un jurado empatado. Pero podría aumentar la calidad de las deliberaciones y, por ende, podrían disminuir los costos del error al requerir que todos estuvieran convencidos de un resultado. En efecto, una regla de unanimidad aumenta el tamaño efectivo del jurado por comparación con la regla de la mayoría. Con todo, requerir una supermayoría o mayoría cualificada, digamos de dos tercios, sería preferible a requerir la unanimidad, debido a que este último requerimiento le da a cada miembro del jurado una capacidad de veto, y existe el riesgo de que uno o dos miembros de un grupo *ad hoc* seleccionado aleatoriamente sean irresponsables, dado que no cargarán con ninguno de los costos de su irresponsabilidad.

El uso de la conscripción para obtener jurados parece ineficiente a primera vista, ya que conduce a una subestimación de los costos sociales del jurado y por consiguiente al uso excesivo de los jurados. Pero sería difícil obtener jurados de personas con antecedentes diversos —jurados presumiblemente más eficientes como aclaradores de los hechos— sin el uso de la compulsión. Un honorario fijado a un nivel que llenara justamente el número de lugares del jurado produciría jurados integrados principalmente por personas de estatus económico bajo. Un honorario fijado a un nivel suficientemente elevado para atraer a personas de ingreso superior crearía una demanda excesiva de los puestos de jurados. Para muchas personas, semejante honorario excedería sus costos de oportunidad; por ello, harían cola en pos de una oportunidad de rendir servicio en el jurado. Y, si los tribunales racionaran la demanda excesiva utilizando algún criterio de capacidad —educación, ocupación o lo que sea—, el jurado perdería su carácter de muestreo al azar, lo que constituye un factor positivo en su capacidad para hacer determinaciones de hechos correctas y pronosticables.

Los jurados no son los únicos legos empleados para adjudicar las disputas legales. La mayoría de los árbitros comerciales no son abogados (sin embargo, la mayoría de los árbitros laborales sí lo son), y, por definición, ninguno de ellos es un juez públicamente empleado. La diferencia entre los árbitros y los jurados es que los primeros se escogen por su pericia en el área en disputa, ¡y los segundos por su falta de pericia! Sin embargo, la diferencia puede

explicarse. Hay un trueque de compensaciones y sacrificios entre la pericia y la imparcialidad. Entre más sepa una persona acerca de algún área de la vida, menos influirán los argumentos y las pruebas de dos rivales sobre su resolución de la disputa; el incremento de su conocimiento es menor. En consecuencia, el experto es más poderoso. Si dos rivales optan (de ordinario antes de la disputa) por someter su disputa a un experto, no hay ninguna razón para no respetar su elección. Pero los jueces y los jurados ejercen el poder coercitivo del gobierno. Ese poder se amortigua al insistir en que no saben tanto acerca de la disputa que sea improbable que escuchen atentamente las presentaciones de los rivales, y se diluye al tener más de un miembro del jurado.[47]

XXII.17. La adjudicación óptima: el caso del contrato

En este capítulo y en el XX hemos pasado revista a una variedad de arreglos institucionales para adjudicar disputas legales. Podemos comprenderlas mejor si consideramos arreglos alternativos para adjudicar una clase particular de disputas, a saber: las disputas contractuales. En un extremo, imaginemos que a un magistrado de profesionalismo cabal, a todas luces competente y honesto, se le asigna determinar el significado de un contrato. Los costos de errores se verían así minimizados. Los costos totales del gobierno podrían ser elevados a causa del salario y otros gastos realizados por un tribunal de alta calidad. Aun así, podrían ser igualmente bajos si la alta calidad de las decisiones del tribunal se tradujera en una tasa de litigación más baja debido a que las partes contractuales tendrían menos incentivos para someter a discusión cuestiones interpretativas de carácter espurio; pues entonces el cuerpo judicial sería más reducido y, por ende, sus gastos serían más bajos. Asimismo, la mayor inversión de recursos gubernamentales por caso se podría compensar con costos de litigación más bajos para las partes en virtud de que los jueces necesitarían menos ayuda de los abogados para alcanzar un resultado correcto. No obstante, la reducción en los costos privados de litigación podría conducir a un incremento en la tasa de litigación. Y la alta calidad del cuerpo judicial también atraería al negocio de resolución de disputas a que se dedican los responsables de hacer arbitraje. Incluso podría hacer que aumentara el número de contratos, una fracción de los cuales habría de dar lugar a disputas litigadas. Los sustitutos de los contratos se vuelven menos atractivos entre más bajos son los costos por la transacción contractual, y es probable que los contratos sean más breves entre más competentes sean los jueces, porque

[47] Por lo que toca a la economía del arbitraje, véanse la siguiente sección de este capítulo y también Orley Ashenfelter, "Arbitrator Behavior", 77 *Am. Econ. Rev. Papers & Proceedings,* 342 (mayo de 1987); William M. Landes y Richard A. Posner, "Adjudication as a Private Good", 8 *J. Legal Stud.,* 235, 236-259 (1979), y XX.1 antes. La economía del juicio por jurados se discute extensamente después en XXIII.3.

los abogados no tendrán que deletrearle cada punto a un intérprete de escaso entendimiento. De este modo, unos costos de transacción contractual más bajos traerán más contratos, lo que puede desembocar en más litigaciones contractuales aun si la tasa de esa litigación desciende. Pero los beneficios sociales netos de los contratos adicionales superarían a los costos representados por litigar el pequeño porcentaje de esos contratos que darían lugar a la litigación. Pues es improbable que el modesto subsidio judicial de los contratos, derivado del hecho de que las partes en una demanda no sobrellevan los gastos en que incurrió la magistratura al exigir el cumplimiento de los contratos, dé por resultado un número mayor de contratos ineficientes.

En el otro extremo, consideremos una magistratura que es incompetente o corrupta, o más probablemente ambas cosas. Los costos que significara el emplear semejante magistratura ya sea para tener claras las intenciones de las partes o para hacer la mejor suposición en lo relativo a la resolución más eficiente de la cuestión interpretativa serían prohibitivos. Las reglas diseñadas para reducir la discrecionalidad judicial, tales como la de requerir que los contratos se formulen por escrito y que siempre sean interpretados de acuerdo con su significado literal, serían eficientes en tales circunstancias. Incrementarían los costos para las partes de la negociación y la redacción, pero los costos totales de los contratos serían probablemente más bajos que si le correspondiera a la magistratura tratar de resolver las ambigüedades. Una magistratura incompetente no podría desempeñar la tarea de modo satisfactorio, y, debido a la incertidumbre en la interpretación, se haría difícil determinar si las decisiones eran corruptas.

Las alternativas de los jueces para resolver las disputas de contratos son los jurados y los árbitros. Una razón común para incluir una cláusula de arbitraje en un contrato es evitar los caprichos de las determinaciones hechas por los jurados, quienes rara vez tienen experiencia comercial; ésa es una razón por la que incluso en los juicios con jurado el significado de un contrato es por lo general abierto a la determinación del juez y no del jurado. En la medida en que los árbitros son considerados intérpretes más confiables que los jueces o los jurados, podemos esperar que los contratos que contengan cláusulas de arbitraje sean más cortos que los contratos que no las tienen.

El motivo para pedir incluir una cláusula de arbitraje en un contrato no es solamente la desconfianza que inspiran quienes, de acuerdo con la convención, se encargan de discernir los hechos; éste incluso puede no ser el motivo principal, considerando cuán a menudo es elegido el arbitraje en sistemas legales que no emplean jurados en los casos civiles. Otros motivos son la privacidad (los registros judiciales son, con raras excepciones, documentos públicos), el deseo que anima a las partes de zanjar sus disputas con base en los usos y costumbres comerciales en vez de en el derecho formal de una jurisdicción particular, la creencia de que los árbitros están menos sujetos a caer presa de ilusiones cognitivas de diferente índole (como el sesgo en retrospec-

tiva) que los miembros del jurado, la disponibilidad de árbitros que pueden dedicar más tiempo a un caso complejo (ya que se les paga por hora y no se les asigna un número fijo de casos) que un juez de instrucción atareado y, como se hizo notar antes en XX.1, la aversión al riesgo, pues se cree que los árbitros tienden a obtener resultados más próximos al punto ubicado a la mitad del camino toda vez que, de proceder de modo distinto, es improbable que salgan elegidos para arbitrajes en el futuro. Adicionalmente, no obstante, la parte contractual que espera ser demandada en lugar de demandar si la relación contractual se viene para abajo puede querer que exista una cláusula de arbitraje en la expectativa de que la propensión de los árbitros a elegir el punto medio habrá de reducir la responsabilidad esperada de la parte. Es más probable que una empresa de corretaje, por ejemplo, sea demandada por un cliente a que ella demande a un cliente. Uno podría pensar que, si la inclusión de una cláusula de arbitraje favoreciera a una parte sobre la otra, la otra parte pondría objeciones o demandaría una compensación. Pero esta última parte podría no molestarse en hacerlo porque el costo esperado de esa acción sería difícil de estimar y sería probablemente ligero habiendo sido descontado por la probabilidad de una ruptura de contrato que condujera a un juicio real. La explicación es paralela a la que da cuenta de que los contratos con forma sesgada hacia un lado en las transacciones del consumidor pudieran constituir un equilibrio competitivo (véase antes IV.9).

Por supuesto, si se tomara literalmente la frase "punto medio del camino", los demandados potenciales no derivarían de ello beneficio alguno. Supongamos que la mitad de los demandantes en alguna clase de casos plantean demandas no meritorias y, por consiguiente, deberían recibir daños de cero, y la otra mitad tienen demandas meritorias y derecho a pago de daños por 10 000 dólares. El demandante no se halla en mejor situación si los árbitros conceden 5 000 dólares en todos los casos. Pero si, como resulta probable, la mayoría de las demandas carecen claramente de méritos, de modo que los árbitros se sienten cómodos al conceder un auxilio de cero en esos casos, su propensión al punto medio del camino sólo operará en los casos meritorios y el efecto consistirá en aminorar la responsabilidad del demandado.

Un indicio de la desconfianza que inspiran los miembros del jurado como intérpretes del contrato lo constituye la práctica de incluir una renuncia al jurado en los contratos. Si el único motivo para el arbitraje es el de evitar a un jurado, tal renuncia resulta una alternativa seductora a una cláusula de arbitraje, especialmente para una parte que no anticipa lograr una ventaja merced al fallo en dirección al punto medio del camino.

XXII.18. Las apelaciones

La parte perdedora de un juicio legal puede apelar a un tribunal superior. La apelación sirve a dos propósitos sociales: reducir los costos del error legal y permitir la creación y el mantenimiento de reglas de derecho uniformes. El primer punto puede parecer inconsistente con el análisis emprendido en una sección previa de la *res judicata* y los impedimentos colaterales. Pero no lo es. El tribunal de apelaciones no es idéntico al tribunal de instrucción, sino que está diseñado especialmente para la función de corregir los errores. De esta manera, la apelación no es una mera repetición del juicio de instrucción de la que no podría esperarse que tuviera un mayor grado de precisión en el fallo.

Si las apelaciones son más probables de ocurrir en el caso de que el tribunal inferior haya cometido un error —y lo son si hay costos fijos no triviales de la apelación (pues entonces no se presentará una apelación a menos que exista una probabilidad significativa de ganar, lo que a su vez implica que hay una probabilidad significativa de error en el tribunal de instrucción)—, la utilización de las apelaciones como un método de corrección del error podría ser superior a la inversión de más recursos en la calidad de los tribunales inferiores a fin de reducir la probabilidad de los errores de esos tribunales. Tales recursos se gastarían en todos los casos, mientras que el costo de una apelación es sufragado sólo en aquellos casos —una fracción moderada del total— que se apelan.[48] Más importante aún, si los tribunales de apelación no estuvieran facultados para corregir los errores, no podrían desempeñar su función sustantiva de elaborar leyes (la producción de precedentes), porque los litigantes no tendrían ningún incentivo para apelar.

En virtud de que de ordinario la parte perdedora no desea mejorar el derecho, parecería que habría muy pocas apelaciones desde un punto de vista social, aun con el moderado subsidio del proceso de apelación (como del de la instrucción) que se proporciona al no cobrar a las partes por los sueldos de los jueces y gastos relacionados. Pero, como vimos en el capítulo anterior, hay un mecanismo equilibrador natural. Si hay demasiado pocas apelaciones en el periodo uno, disminuirá la producción de precedentes por los tribunales de apelación, lo que dificultará que los individuos arreglen sus disputas (porque no pueden ponerse de acuerdo sobre cómo se resolverán probablemente las disputas en el tribunal), conduciendo a un aumento de la litigación y, por ende, a más apelaciones en el periodo dos.

Los tribunales de apelación hacen una revisión plenaria de cuestiones de derecho puro; es decir, no otorgan ninguna deferencia a las opiniones de los jueces de instrucción sobre tales cuestiones. Si lo hicieran, el derecho variaría

[48] Steven Shavell, "The Appeals Process as Means of Error Correction", 24 *J. Legal Stud.,* 379 (1995).

de un juez de instrucción a otro de la misma índole. Pero los tribunales de apelación sí otorgan una deferencia muy amplia a la determinación de los hechos del juez de instrucción o el jurado. Quien determina los hechos invertirá usualmente mucho más tiempo en un caso que se va al tribunal de instrucción que los jueces de apelación, y obtendrá una mejor captación intuitiva de los hechos. Y, dado que los hechos son diferentes de un caso a otro, la coherencia de la ley no requiere que todos los casos se decidan de la misma forma, sino que todos los jueces apliquen las mismas reglas.

Los tribunales de apelación no revocan los errores inocuos. Por definición, es improbable que tal error conduzca a un resultado diferente si el caso vuelve a instruirse. El beneficio esperado de la revocación en tal caso es bajo en relación con el costo de los nuevos procedimientos que seguirían en el tribunal de instrucción si se revocara el fallo.[49]

Una interrogante de importancia acerca de la posibilidad de apelación es ésta: ¿cuándo podrá apelarse el fallo del tribunal de instrucción? ¿Cuando se promueve el juicio o hasta el final del procedimiento en el tribunal de instrucción?

La presunción favorece la postergación de la revisión hasta el final del procedimiento en el tribunal de instrucción. Pero cuando se permiten las apelaciones inmediatas de los fallos intermedios (interlocutorios) se obtienen dos beneficios. El primero es evitar la demora en determinar lo correcto del fallo. El segundo es eliminar procedimientos dilatados en el tribunal de instrucción si, por ejemplo, este tribunal negó erróneamente una moción para que se deseche la reclamación. Pero al lado de estos dos beneficios hay dos costos. El primero es que el procedimiento en conjunto podría demorarse por interrupciones frecuentes para las apelaciones de los fallos interlocutorios. El segundo es que la posposición de la revisión de apelación podría eliminar procedimientos de *apelación* innecesarios, porque muchos de los fallos de los jueces de instrucción en el curso de un proceso dejan de tener importancia para el momento en que termina el proceso. Supongamos, por ejemplo, que un juez falla en contra del demandante una y otra vez (en cuestiones de admisión de las pruebas, por ejemplo), pero al final el demandante gana de todos modos. En un sistema que permitiera libremente las apelaciones interlocutorias, el demandante podría haber apelado cada uno de esos fallos en el momento en que se emitieran, pero si no se permitieran las apelaciones interlocutorias todos esos fallos dejarían de tener importancia para el fallo final.

Además, en lugar de tener que considerar 10 apelaciones en el mismo juicio, el tribunal de apelaciones sólo tiene que considerar una apelación con (quizá) 10 cuestiones; y, en la medida en que las cuestiones se relacionen entre sí —por ejemplo, porque se basan en los mismos hechos—, la apelación única podría consumir menos tiempo de los jueces de apelación, aunque

[49] El error inocuo se discute más extensamente después en XXIII.5.

tenga muchas cuestiones, que 10 apelaciones de una sola cuestión en el mismo juicio.

Un régimen de apelaciones interlocutorias economiza el tiempo de los tribunales de instrucción a costa de los tribunales de apelación, mientras que una regla de fallo final hace lo contrario. Esta última es más eficiente si el tiempo del juez de apelación constituye una oferta más escasa que la del tiempo del juez de instrucción. Lo es. El número de los jueces de apelación en un sistema unitario no puede expandirse indefinidamente sin añadir nuevos estratos de tribunales de apelación, ya que una tarea vital del proceso de apelación —el mantenimiento de una uniformidad y una congruencia razonables del derecho— no puede realizarse con eficacia si hay más de un puñado de jueces. Si la carga de trabajo de las apelaciones es tal que se necesitan más de unos cuantos jueces para desahogarla, estos jueces deberán dividirse en colegios más pequeños, pero entonces surge un problema de coordinación de los colegios. Finalmente, la única respuesta es otro estrato de revisión de apelación. El sistema federal tiene ya un sistema de tres estratos, aunque está limitado el acceso al tercer estrato, la Corte Suprema. No sería viable un cuarto estrato.

Adviértase el contraste con otras organizaciones, tales como las empresas de negocios, donde es común que exista una jerarquía de múltiples niveles. Los empleados colocados en el nivel inferior manejan asuntos de rutina. Se opta por girar una cuestión que no es rutinaria uno o dos o muchos estratos administrativos hacia arriba, dependiendo de su dificultad (a esto se le llama "administración por excepción"). Esto no funciona para los sistemas judiciales, lo que equivale a decir que estos sistemas presentan pronunciadas deseconomías de escala. ¿Qué hacer, entonces, a fin de dar cabida a la demanda acrecentada? Una posibilidad es la de los tribunales especializados; pues puede haber tantos tribunales especializados como haya especialidades y, así, el tamaño medio de cada tribunal será más pequeño (el negocio judicial se divide entre un número más grande de tribunales) y los costos de transacción (que se incrementan con el tamaño de un tribunal de apelaciones) son más bajos. Otra posibilidad es la de delegar las responsabilidades adjudicatarias en agencias administrativas, según vimos en el capítulo precedente, sujetas a la revisión judicial (la alternativa a limitar la revisión de apelaciones a las apelaciones dentro de la agencia en lugar de a aquellas que van de la agencia al tribunal es objetable porque los costos de errores serían altos, dado que una agencia goza de menos independencia política que un tribunal). Una posibilidad más es la de incrementar la productividad de cada estrato administrativo al expandir el personal, digamos, añadiendo asistentes jurídicos y abogados del personal y funcionarios judiciales subordinados (tales como los jueces magistrados). Según señalamos previamente, ésta ha sido la respuesta fundamental de los tribunales estadunidenses al incremento en el número de casos.

A pesar de la renuencia de los tribunales de apelación a escuchar múltiples apelaciones del mismo caso, esto resulta en algunos casos inevitable. La de-

mora en conseguir la revisión de un fallo interlocutorio podría ser muy costosa cuando el fallo fuese una orden preliminar que obligue al demandado a cerrar su negocio, o una disposición sobre una cuestión potencialmente final que si se decidiera en sentido contrario eliminaría un procedimiento de un año de duración. Por lo tanto, no es sorprendente que una ley permita que los fallos interlocutorios que otorgan o niegan órdenes sean apelados de inmediato (véase antes XXII.3). Y la doctrina de la "orden colateral" permite que las órdenes interlocutorias sean apeladas de inmediato si implican cuestiones separadas de los méritos de la litigación y una apelación inmediata es necesaria para impedir un daño irreparable contra el apelante. Supongamos que el tribunal de instrucción se niega a hacer que el demandado otorgue una fianza para asegurar que si pierde el juicio será capaz de efectuar el rembolso de ciertos gastos del demandante al que éste tendría derecho. Si la orden no es apelable de inmediato y el demandante gana y resulta que el demandado no tiene dinero, el demandante se quedará con un palmo de narices. Y, dado que el otorgamiento de la fianza está separado de los méritos del caso, no habrá deseconomía judicial por el hecho de tener que saber de apelaciones separadas sobre la fianza y sobre los méritos del caso.

Aun así, es patente que el fuerte interés social en la irrevocabilidad de la litigación (véase antes VIII.6) requiere de una fuerte presunción a favor de limitar las apelaciones a los fallos finales.

XXII.19. La elección de la ley

Supongamos que un residente del estado A, mientras conduce en el estado B, lesiona a un residente de B, quien lo demanda.[50] ¿Cuál es la ley estatal que deberá utilizarse para decidir los derechos de las partes: la de A o la de B? Ambos estados tienen un interés en el resultado del juicio. Al nivel más simple, A se beneficiará si su residente gana, porque éste tendrá más dinero, y B se beneficiará si su residente gana. Estos beneficios se contrarrestan y pueden omitirse. Pero A tiene también un interés en que sus residentes puedan conducir en el estado B sin restricciones indebidas, y B tiene un interés en proteger a su residente para que no sea lesionado por un conductor negligente. En otras palabras, ambos estados tienen en la disputa un interés de asignación y uno de distribución. Y aquí resulta significativo que el accidente ocurriera en B. Presumiblemente, las reglas de B se ajustan a las condiciones de la conducción —el estado de los caminos, el clima, etc.— existentes en B. Dado que B tiene así una ventaja reguladora comparativa[51] (sujeta a una calificación que examinaremos pronto) en lo tocante a los accidentes

[50] Discutida más extensamente adelante en XXVII.10.

[51] La elección del criterio de la ley propuesta en William F. Baxter, "Choice of Law and the Federal System", 16 *Stan. L. Rev.,* 1 (1963).

que ocurren en B, hay un argumento económico en favor de la regla tradicional del derecho común de *lex loci delicti*, esto es, de que la ley del lugar donde ocurrió el cuasidelito es la ley que se aplicará en un juicio por cuasidelitos, dondequiera que se promueva el juicio para remediar ese cuasidelito.

Pero esta ley no funciona bien en todos los casos. Se plantean sutiles dificultades en un caso donde dos residentes del estado A se ven implicados en una colisión en el estado B. Las reglas del cuasidelito de B se adaptarán mejor a los factores específicos de la ubicación, tales como el estado de los caminos y las condiciones climáticas, pero las reglas de los cuasidelitos de A se adaptarán mejor a los factores específicos de las personas, tales como la capacidad relativa para tener cuidado o la cantidad de daños necesaria para compensar plenamente a la víctima del accidente, y pueden ser los únicos factores relevantes para el caso en particular. Y está además la cuestión de dónde ocurrió precisamente el cuasidelito. Técnicamente, ocurre en el estado donde tuvo lugar la lesión, porque no existe un cuasidelito de no haber una lesión o perjuicio; pero ¿no podría pensarse que el estado donde sucedió el acto de negligencia posee la misma ventaja regulatoria comparativa, en lugar de un estado diferente, en el que se diera el caso de que sobreviniera la lesión? ¿Tiene esto sentido, empero, en un caso de responsabilidad de productos considerado separadamente de un típico caso de accidente automovilístico? En la mayoría de los casos cuasidelictivos, la víctima y el victimario son residentes del mismo estado, y, presumiblemente, la ley de cuasidelitos del estado logra mantener un equilibrio entre los intereses de las víctimas potenciales y los de los victimarios potenciales. Con todo, este balance se halla ausente en muchos casos de responsabilidad por productos debido a que un estado puede ser un gran fabricante de productos peligrosos que se consumen principalmente en otra parte.

En un caso de difamación en los medios a escala nacional, en el que el perjuicio causado a la reputación del demandante puede haber ocurrido en 50 estados distintos, debería aplicarse el estatuto de limitaciones que rige en el estado del victimario a fin de reducir la compra de foros legales (o búsqueda de tribunales más propensos a emitir fallos favorables). Esto es, el mismo cuerpo legal debiera aplicarse dondequiera que se levante la demanda, a condición de que no exista una diferencia sustancial entre los sistemas de tribunales de las jurisdicciones donde pueda establecerse el juicio.[52]

Complicaciones de esta laya han empujado a la mayoría de los estados a adoptar un enfoque más complejo de la selección de asuntos legales en casos de cuasidelitos que el de *lex loci delicti* (aunque éste continúa siendo el enfoque por presunción) al esforzarse en determinar, según una base de caso por caso, cuál estado afectado por la demanda tiene un "interés" más grande en aplicar su ley. Desde un punto de vista económico, la controversia no debiera

[52] Kamelgard *vs.* Macura, 585 F.3d, 334 (7th Cir., 2009).

ser la de los intereses, sino la ley de cuál estado "encaja" mejor en las circunstancias de la disputa. Supongamos que un asunto clave es el de decidir cuál ley de limitaciones se aplicará, si la de este o aquel estado. Si el propósito de la ley de limitaciones es la disminución de los costos del error asociados con el uso de pruebas viejas, hay un sólido argumento en favor de la aplicación de la ley del estado donde se instruye el caso porque presumiblemente refleja la competencia de los tribunales de ese estado para tratar las pruebas viejas. (Ésta es una razón de que los tribunales apliquen sus propias reglas procedimentales incluso cuando están aplicando la ley sustantiva de otra jurisdicción.) Pero si el propósito de la ley de limitaciones es el de permitir que los individuos planeen sus actividades con mayor certeza, hay un argumento en favor de la aplicación de la ley de limitaciones del estado del victimario, porque es éste quien está sujeto a la incertidumbre. Si los estatutos de limitaciones de ambos estados tienen ambos propósitos, entonces el estatuto de limitaciones vigente en el estado del victimario es el que debiera aplicarse, siempre y cuando no haya una diferencia sustancial entre los sistemas de tribunales de las jurisdicciones donde pueda establecerse el juicio.

O supongamos que la disputa en un juicio por incumplimiento de contrato entre residentes de estados diferentes es la capacidad del promitente para hacer una promesa vinculante (por ejemplo, ¿deberá tener 21 años de edad o con 18 es suficiente?); las reglas del estado de residencia del promitente tendrían la ventaja comparativa al resolver la disputa, porque presumiblemente estas reglas se basan en las capacidades promedio de los residentes de ese estado.

La tendencia de los tribunales a sustituir las reglas tradicionales del derecho común —más simples, y entre las que se cuentan las relativas al lugar del cuasidelito o el lugar de elaboración del contrato— por criterios multifactoriales de elección de la ley ha aumentado la proporción de casos en los que el tribunal aplica la ley de su propio estado. Pues ésa sería la tendencia natural de un tribunal, y se trata de una preferencia que resulta más fácil de complacer cuando los criterios de selección, siendo múltiples, son vagos. Posee ventajas de eficiencia: los jueces están más familiarizados con la ley de su propio estado que con la de otra jurisdicción; la mayoría de los casos son litigados por abogados locales que, por consiguiente, están más al tanto de la ley del estado donde ocurre el juicio que con la de otros estados; y las decisiones basadas en la ley local incrementan el acervo de precedentes, puesto que la decisión de un tribunal sobre la interpretación de la ley de otro estado tendrá solamente un significado débil como precedente en ese estado.[53]

Por lo general, los tribunales exigirán el cumplimiento de las disposiciones en materia de elección de la ley contractual. Un tribunal debiera rehusarse a hacer cumplirla (excepto en casos de fraude) sólo si la ley elegida es

[53] Véase Stuart E. Thiel, "Choice of Law and the Home-Court Advantage: Evidence", 2 *Am. Law & Econ. Rev.,* 291 (2000).

de tal índole que el tribunal tiene dificultades para aplicarla, tal vez debido a que la ley procede de una nación extranjera. Empero, no debería exigírseles a las partes de los contratos que incluyeran disposiciones sobre elección de la ley. Semejante requisito incrementaría el costo de llevar a cabo la contratación. Dado que son escasas las disputas contractuales que terminan en el tribunal, resulta mejor rellenar una laguna en el contrato dentro de un caso litigado que obligar a las partes a negociar en torno a un asunto específico, tal como el de la elección de la ley.

XXII.20. La economía de la profesión de abogado

El papel más conspicuo y socialmente problemático del abogado se encuentra en la instrucción de los juicios; por lo tanto, este capítulo parece tan bueno como cualquier otro para echar una ojeada a la economía de la profesión misma y para cuestionarnos si los abogados aumentan o disminuyen el producto social. Por supuesto, los abogados hacen otras cosas aparte de litigar los casos; sobre todo asesoran a las personas que están considerando un curso de acción que podría generar conflictos legales. El consejo de los abogados ayuda a evitar violaciones inadvertidas de la ley, y en esa medida reduce la litigación y (probablemente) promueve el bienestar social. No obstante, al mismo tiempo, impide también cierto cumplimiento inadvertido (pero no por ello menos valioso socialmente) con la ley. El cliente podría aprender del abogado que es improbable que un curso de acción que, según conjeturaba él correctamente, era ilegal genere una sanción grave. O el abogado podría conocer las fisuras y lagunas que permiten al cliente una ligera restructuración del curso de acción propuesto a fin de cumplir con la letra y violar el espíritu de la ley. Una de las formas del consejo legal económicamente más dudosas es el consejo acerca de cómo minimizar nuestra responsabilidad fiscal. Una disminución de los impuestos de un contribuyente tenderá a verse contrarrestada a largo plazo por un aumento de los impuestos de otro contribuyente, de modo que, a no ser que los impuestos del segundo contribuyente tengan menos efectos de asignación o distribución indeseables —y en promedio no los tendrán—, el gasto del primer contribuyente en asesoría legal será un costo social de peso muerto.[54]

Este ejemplo ilustra el punto de que el papel del abogado se guía por los beneficios privados antes que por sus beneficios sociales. Esto es igualmente cierto en el rubro de la litigación. El privilegio abogado-cliente (analizado en el siguiente capítulo) le prohíbe a aquél revelar sin permiso del cliente la información que haya recibido de él, aunque esa información demuestre que

[54] Louis Kaplow, "Accuracy, Complexity, and the Income Tax", 14 *J. Law, Econ. & Organization,* 61 (1998).

la reclamación o la defensa del cliente no tienen mérito. Esto ilustra el hecho de que el abogado no tiene (en general) ninguna obligación de revelar al tribunal la información que sea dañina para la causa de su cliente; la información no tiene que haber provenido del cliente, de modo que no cae necesariamente en el campo del privilegio abogado-cliente.

Los Estados Unidos cuentan con un millón de abogados. ¿Son demasiados? Un estudio realizado por distinguidos economistas descubrió, tras corregir por otros factores, que el crecimiento económico de una nación se correlaciona positivamente con el número de ingenieros y negativamente con el número de abogados.[55] La explicación de los autores es que los ingenieros producen riqueza al diseñar métodos para la reducción de los costos o el mejoramiento de los productos, mientras que los abogados simplemente producen redistribuciones de la riqueza (como en el ejemplo de los impuestos). Pero sabemos que la protección de los derechos de propiedad es fundamental para la creación de riqueza, y los abogados desempeñan un papel importante en esa protección. También contribuyen a la producción fuera del mercado, lo que se omite en las medidas habituales del crecimiento económico, como el aire y el agua limpios; el control de la contaminación es una actividad muy intensiva en abogados. Otra forma importante de la producción fuera del mercado, y de nuevo intensiva en abogados, es la protección de las libertades civiles. Como los bomberos y las compañías de seguros, gran parte de lo que proveen los abogados son opciones que son valiosas aunque nunca se ejerzan: crean la tranquilidad mental que acompaña al conocimiento de que se dispone de una maquinaria para el respeto de nuestros derechos. Por supuesto, desde el punto de vista de los demandados civiles potenciales (y de las víctimas del delito), éstas podrían ser opciones negativas… pero no por completo. El hecho de estar sujeta a una obligación legalmente exigible podría aumentar la utilidad de una persona al permitirle participar en transacciones ventajosas que de otro modo le estarían vedadas. Si los acreedores no tuvieran remedios legales contra los deudores, a la gente le resultaría difícil obtener préstamos o, como mínimo, tendrían que pagar tasas de interés muy elevadas. En consecuencia, los abogados que ayudan a los acreedores a ejercitar sus remedios proveen un beneficio indirecto a los deudores.

Tradicionalmente, el derecho era, y en un grado disminuido aún lo es, una profesión. El derecho y la medicina son las dos profesiones en las que uno piensa primero, las profesiones en las que parece cristalizar de manera sumamente perfecta el concepto de la profesión como encarnación de la habilidad del experto. Entre los profesionistas están incluidos, aparte de los abogados y los médicos, los arquitectos, los enfermeros, los maestros, los ingenieros, el clero y los oficiales militares (la lista no es exhaustiva), y los que nos ocupan cons-

[55] Kevin M. Murphy, Andrei Schleifer y Robert W. Vishny, "The Allocation of Talent: Implications for Growth", 106 *Q. J. Econ.,* 503 (1991).

tituyen un importante segmento de la economía. Ya mencioné que hay un millón de abogados; existen más de 700 000 médicos.

En el modelo de los negocios, el objetivo es la ganancia en un ambiente competitivo que funciona bajo un esquema darwiniano; el riesgo es omnipresente y tanto las ganancias extraordinarias como las pérdidas devastadoras son posibilidades reales. En un ambiente tal, el empleo y el liderazgo atraen a muchos y repelen a muchos. Las personas que se sienten atraídas tienden a ser agresivas y atrevidas; las que experimentan repulsión tienden a ser cautelosas y reflexivas. En el modelo profesional tradicional, el riesgo que se da tanto en sentido ascendente como descendente es eliminado mediante una combinación de regulación y de ética, y ambas están dirigidas a refrenar la competencia. Una vez logrado esto, el abogado o el médico pueden aspirar a recibir un ingreso cómodo propio de la clase media alta, si bien es improbable que se vuelvan ricos. El resultado, en combinación con la exigencia de hacer estudios de posgrado y presentar los exámenes de calificación para el ingreso a la profesión y la práctica de someter a los aspirantes a la disciplina profesional, consiste en atraer a un tipo de persona muy diferente al empresarial: una persona más estudiosa, más intelectual, con mayor aversión al riesgo.

¿Por qué valora la sociedad a esas personas y crea un nicho cómodo para ellas? Algunos bienes y servicios son tan complejos que dificultan a los consumidores evaluar la calidad de esas ofertas; y cuando esto ocurre el proveedor tiene que ser regulado, ya sea directamente como en el caso de la regulación de la industria farmacéutica realizada por la Administración de Alimentos y Medicamentos o indirectamente como en el modelo profesional, en el que las condiciones para convertirse en miembro de una profesión fomentan la selección individual por personas que tienen todas las probabilidades de ser confiables, responsables y éticas debido a que están menos inclinadas a ahorrar tiempo y esfuerzo para fraguar un asesinato.

Hasta fecha bastante reciente, la profesión de abogado se conformaba muy de cerca al modelo profesional, al menos al nivel de la empresa de grandes dimensiones. Un cuestionamiento temprano y muy parcial de esta pauta se hizo visible con el advenimiento de la facturación por hora, que había sido poco común antes de la segunda Guerra Mundial. Hasta entonces, un despacho de abogados generalmente facturaba por servicios prestados sin especificar ítems. Pero ésa es una práctica viable sólo cuando los clientes depositan una gran cantidad de confianza en sus abogados, yendo en consonancia con el modelo legal. La facturación por hora requiere el señalamiento de ítems, y esto le brinda al cliente alguna base para evaluar el costo en que incurrió el abogado. Con todo, la facturación por hora conlleva fijar el precio igual al costo más un margen de ganancia, lo cual no es eficiente. Incita a usar material de relleno, por ejemplo, bajo la forma de establecer nuevos precios del tiempo utilizado por nuevos asociados que carecen de experiencia.

En derecho, el modelo profesional empezó a declinar en el decenio de 1970 con el comienzo del movimiento de desregulación, el cual relajó las restricciones a la competencia en los servicios legales. La tendencia se mantuvo en las décadas subsiguientes y estuvo marcada por un aumento de la diseminación de las ganancias dentro de los despachos jurídicos, un aumento de la dispersión en las dimensiones de las firmas de abogados y un incremento en la rotación —en particular, la pauta seguida por los abogados exitosos de desplazarse de firma en firma (llevándose consigo a sus clientes) en búsqueda de ingresos más elevados—. La tendencia se aceleró después del colapso financiero de septiembre de 2008; una respuesta de las corporaciones a la adversidad fue la de presionar a las firmas de abogados en pos de tasas más bajas. Un resultado de ello fue la consolidación de las firmas. Muchas de éstas, ubicadas en la delantera, constituyen hoy en día negocios globales, algunos con recaudaciones anuales que exceden los mil millones de dólares. Administradas de manera profesional, son muy competitivas, con lo que ganan y pierden clientes, ganan y pierden abogados del nivel máximo capaces de llevarse consigo a los clientes y tienen bajo la mano las riendas del ciclo económico, lo que trae por consecuencia, en el plano inferior, los despidos irrevocables, los despidos temporales y los recortes o congelamientos salariales que se han atestiguado en los años recientes. Las horas de Nueva York han devenido la norma por doquier.

En la actualidad, los abogados corporativos no se conforman con un ingreso característico de la clase media alta. No pueden permitirse el lujo de resignarse a ello, en virtud del incremento habido en el riesgo que enfrentan. Ya quedan pocas firmas de abogados que sigan practicando la compensación "en fila india", en la que a todos los socios provenientes de la misma época o generación se les da la misma remuneración: un método de compensación para minimizar el riesgo que fungía como norma en las firmas de grandes proporciones. Hoy en día un abogado encara el riesgo, si su productividad declina, de ver cómo declina a la par su ingreso, o incluso de ser sacado a empujones de la firma; y, para amortiguar ese riesgo, desea ganar tanto como pueda mientras esté en su capacidad hacerlo.

Idos son los tiempos de la confianza inspirada en el cliente. Están por irse los de la facturación por hora. La fijación del precio igual al costo más un margen de ganancia no resulta eficiente; lo es la puja competitiva de precios fijos con ajuste, ya sea que esté especificado en la puja o que sea negociado, a los casos en que ha habido un exceso en el cumplimiento de los objetivos de desempeño o no se han cumplido del todo. Está próxima a irse, con toda probabilidad, la proscripción de la venta de intereses accionarios en las firmas de abogados a quienes no son abogados, lo que ha forzado a las firmas a recaudar el capital que necesitan principalmente mediante el préstamo; así se han inhibido el crecimiento, las ganancias y la administración profesional.

Una vez que se relajan los límites legales y éticos impuestos a la competencia entre abogados, la índole riesgosa que de forma subyacente presenta la

práctica de las firmas se manifiesta en toda su potencia, libre de ataduras. A semejanza de los bancos, si bien de manera menos dramática, el capital de las firmas de abogados es a corto plazo, pero sus activos son a largo plazo. El capital principal es el capital humano —los socios más exitosos y sus clientes—, y ese capital es de corto plazo; un socio puede dejar la firma, acompañado de sus clientes, dejando apenas aviso o ni siquiera eso. Esto puede provocar una estampida en la firma, como ocurrió en el colapso de la empresa Dewey LeBoeuf. Los activos de una firma de abogados, incluidos sus cuentas pendientes de cobro y sus clientes futuros (correspondientes a los préstamos de un banco), no pueden ser liquidados en cualquier momento a fin de que los activos de la firma se ajusten a su capital en declive y a su deuda fija. La mayor competitividad de la industria de servicios legales posee implicaciones para la ética legal. Los abogados tienen dos clases de obligaciones éticas: *1)* con el cliente (por ejemplo, evitar el cobro excesivo y los conflictos de interés) y *2)* con la sociedad (por ejemplo, no sobornar el perjurio). La competencia *podría* perjudicar al primer tipo de obligación ética; entre menos rentable sea (debido a su mayor condición competitiva) la práctica del derecho, menor será el castigo financiero por ser expulsado de la barra como una sanción por la violación de la obligación ética para con nuestro cliente. Pero en general no esperamos que los vendedores competitivos sean menos honestos con sus clientes que los monopolistas. Sin embargo, es virtualmente seguro que la competencia habrá de minar la obligación ética del abogado para con la sociedad, ya que, entre más competitivo sea un mercado, será menos probable que los vendedores sacrifiquen los intereses de sus clientes en aras de un beneficio social por el que no se compensa a los vendedores.

Lecturas recomendadas

Baird, Douglas G., Robert H. Gertner y Randal C. Picker, *Game Theory and the Law,* cap. 8 (1994).

Bar-Gill, Oren, y Omri Ben-Shahar, "The Prisoner's (Plea Bargain) Dilemma", 1 *J. Legal Analysis,* 737 (2009).

Bebchuk, Lucian Arye, "A New Theory Concerning the Credibility and Success of Threats to Sue", 25 *J. Legal Stud.,* 1 (1996).

——— y Howard Chang, "An Analysis of Fee Shifting Based on the Margin of Victory: On Frivolous Suits, Meritorious Suits, and the Role of Rule 11", *25 J. Legal Stud.,* 371 (1996).

Cooter, Robert D., y Daniel L. Rubinfeld, "Economic Analysis of Legal Disputes and Their Resolution", 27 *J. Econ. Lit.,* 1067 (1989).

Criminal Law and Economics, caps. 5-7, 14 (Nuno Garoupa [comp.], 2009).

Danzon, Patricia Munch, y Lee A. Lillard, "Settlement Out of Court: The Disposition of Medical Malpractice Claims", 12 *J. Legal Stud.,* 345 (1983).

Easterbrook, Frank H., "Criminal Procedure as a Market System", 12 *J. Legal Stud.,* 289 (1983).

Economics of Commercial Arbitration and Dispute Resolution (Orley C. Ashenfelter y Radha K. Iyengar [comps.], 2009).

Economics of Conflict of Laws (Erin O'Hara [comp.], 2007).

Fon, Vincy y Francesco Parisi, "Litigation and the Evolution of Legal Remedies: A Dynamic Model", 116 *Pub. Choice,* 419 (2003).

Guzman, Andrew T., "Choice of Law: New Foundations", 90 *Geo. L. J.,* 883 (2002).

Hay, Bruce L., "Fee Awards and Optimal Deterrence", 71 *Chi.-Kent L. Rev.,* 505 (1995).

Jensen, Michael C., William H. Meckling y Clifford G. Holderness, "Analysis of Alternative Standing Doctrines", 6 *Int'l Rev. Law & Econ.,* 205 (1986).

Kaplow, Louis, "On the Optimal Burden of Proof", 119 *J. Pol. Econ*., 1104 (2011).

Kornhauser, Lewis A., "Judicial Organization and Administration", en *Encyclopedia of Law and Economics,* vol. 5: *The Economics of Crime and Litigation,* 27 (Boudewijn Bouckaert y Gerrit De Geest [comps.], 2000).

Kramer, Larry, "Rethinking Choice of Law", 90 *Colum. L. Rev.,* 277 (1990).

Landes, William M., "An Economic Analysis of the Courts", 14 *J. Law & Econ.,* 61 (1971).

———, "The Bail System: An Economic Approach", 2 J. Legal Stud., 79 (1973).

——— y Richard A. Posner, "The Economics of Anticipatory Adjudication", 23 *J. Legal Stud.,* 683 (1994).

Lederman, Leandra, "Which Cases Go to Trial?: An Empirical Study of Predictors of Failure to Settle", 49 *Case W. Res. L. Rev.,* 315 (1999).

Miceli, Thomas J., "Do Contingent Fees Promote Excessive Litigation?", 23 *J. Legal Stud.,* 211 (1994).

Mitchell Polinsky, A. y Daniel L. Rubinfeld, "Sanctioning Frivolous Suits: An Economic Analysis", 82 *Geo. L. J.,* 397 (1993).

Priest, George L., "Private Litigants and the Court Congestion Problem", 69 *B. U. L. Rev.,* 527 (1989).

Procedural Law and Economics (Chris William Sanchirico [comp.], 2012).

Ribstein, Larry E., "Ethical Rules, Agency Costs, and Law Firm Structure", 84 *Va. L. Rev.,* 1707 (1998).

Rosen, Sherwin, "The Market for Lawyers", 35 *J. Law & Econ.,* 215 (1992).

Shavell, Steven, "Suit, Settlement, and Trial: A Theoretical Analysis under Alternative Methods for the Allocation of Legal Costs", 11 *J. Legal Stud.,* 55 (1982).

———, "Legal Advice about Contemplated Acts: The Decision to Obtain Advice, Its Social Desirability, and Protection of Confidentiality", 17 *J. Legal Stud.,* 123 (1988).

———, "The Appeals Process as a Means of Error Correction", 24 *J. Legal Stud.,* 379 (1995).

Spurr, Stephen J., "An Economic Analysis of Collateral Estoppel", 11 *Int'l Rev. Law & Econ.*, 47 (1991).
The Evolution of Efficient Common Law (Paul H. Rubin [comp.], 2007).
Thiel, Stuart E., "Choice of Law and the Home-Court Advantage: Evidence", 2 *Am. Law & Econ. Rev.*, 291 (2000).
Weisburst, Sanford I., "Judicial Review of Settlements and Consent Decrees: An Economic Analysis", 28 *J. Legal Stud.*, 55 (1999).
Whincop, Michael J. y Mary Keyes, *Policy and Pragmatism in the Conflict of Laws* (2001).

Problemas

1. ¿Puede imaginar una justificación económica para remplazar la demanda simplificada con la exposición de alegatos o alegato de emisión, como se hace en las Reglas Federales del Procedimiento Civil?

2. ¿Debiera una de las partes de un juicio estar facultada para exigir la presentación de un testigo?

3. En Clauss *vs.* Danker, 264 D. Supp. 246 (S.D.N.Y., 1967), una acción de cuasidelito, el demandante buscaba el descubrimiento de los detalles acerca de la póliza de seguro de responsabilidad del demandado, aunque tales detalles no serían admisibles como prueba en la instrucción del caso. El tribunal dijo que sería conveniente permitir tal descubrimiento, sobre todo para facilitar el arreglo, pero que de acuerdo con las Reglas Federales del Procedimiento Civil no podría ordenarse el descubrimiento de la cobertura del seguro. En 1970 se enmendaron las reglas para permitir el descubrimiento. El comité asesor observó: "La revelación de la cobertura del seguro permitirá que los abogados de ambas partes hagan la misma evaluación realista del caso, de modo que el arreglo y la estrategia de la litigación se basen en el conocimiento y no en la especulación. Ello conduciría al arreglo y evitaría una litigación prolongada en algunos casos, aunque en otros podría tener un efecto contrario".

¿Es probable que el cambio de las reglas aumente o disminuya la frecuencia del arreglo? ¿Cómo difiere el análisis, si acaso, del de la Regla 35 discutida antes en XXII.6?

4. Compare los costos de la reducción del delito mediante: *a)* la disminución de la carga de la prueba para el fiscal, y *b)* el hecho de volver más específicas las prohibiciones delictivas.

5. ¿Cuáles serían las consecuencias si, en lugar de exigir que el acusado otorgara una fianza, se requiriera que el Estado pagara al acusado los costos en los que incurrió por permanecer en prisión durante el juicio?

6. ¿Sería una buena idea, desde un punto de vista económico, que se compensara al acusado que es absuelto de una acusación penal? ¿Cuál sería la medida de compensación apropiada? ¿Cuáles implicaciones tendría la compensación, si acaso, para el estándar de la prueba apropiada en un caso penal?

7. En un juicio de daños triples según las leyes antimonopólicas, ¿debería exigirse al demandante que establezca la responsabilidad del demandado fuera de toda duda razonable?

8. Dos preguntas acerca de la indemnización de los honorarios del abogado de la parte ganadora:

a) Supongamos que la parte ganadora estaba representada por una empresa legal de interés público. ¿Debieran basarse los honorarios del abogado en los costos efectivamente sufragados por la empresa, o en lo que abogados de igual competencia habrían cobrado por los servicios prestados por la empresa? Véase Blum *vs.* Stenson, 465 U.S., 886 (1984).

b) ¿Debiera ser un multiplicador del riesgo un componente de los honorarios "razonables" de los abogados? Supongamos que el demandante tenía sólo una probabilidad de 20% de ganar y que, si hubiera perdido, su abogado no habría obtenido nada porque había convenido en tomar el caso por un honorario contingente. Si el demandante gana, ¿debiera multiplicarse el honorario del abogado por cinco a fin de compensarlo por el riesgo que corrió? Véanse Laffey *vs.* Northwest Airlines, Inc., 746 F.2d, 4 (D.C. Cir., 1984); McKinnon *vs.* City of Berwyn, 750 F.2d, 1383 (7th Cir., 1984).

9. Uno de los varios beneficiarios de un fideicomiso contrata a un abogado para demandar al administrador fiduciario. Si derivado del juicio aumentan los activos del fideicomiso, la ley permitirá que el abogado reclame una parte de sus honorarios a los otros beneficiarios. ¿Tiene esto sentido económico? Véase Saul Levmore, "Explaining Restitution", 71 *Va. L. Rev.,* 1 (1985).

10. A fin de obtener una orden preliminar, un demandante debe demostrar, como hemos visto, que sufrirá un daño irreparable si se niega la orden. Desde un punto de vista económico, ¿cuál de los siguientes tipos de daño presuntamente irreparable debiera ser una base para una orden preliminar en un juicio de daños?

a) El demandante podría volverse insolvente antes de la terminación del juicio.

b) El demandante podría declararse en quiebra antes de la terminación del juicio como resultado de la presunta infracción del demandado.

c) Los daños serán difíciles de calcular.

11. ¿Debiera otorgarse un efecto de impedimento colateral a un fallo que no se apeló?

12. ¿Esperaría que una cláusula contractual que estipule el arbitraje de las disputas surgidas a partir del contrato fuese más común en los contratos de forma estándar impresos o en los contratos individualmente negociados?

13. Desde el punto de vista de la teoría de juegos, compare el enfoque convencional del desacato civil, que permite a un juez ordenar que un testigo que desafía una orden de testificar sea encarcelado hasta que testifique o convenza al juez de que no testificará por mucho que dure el encarcelamiento, con una regla que fije un término de prisión definido por el desacato. Véase Linda S. Beres, "Games Civil Contemnors Play", 18 *Harv. J. Law & Pub. Policy,* 795 (1995).

14. Evalúe la proposición siguiente: La importancia de la *res judicata* no reside en que impida la relitigación, ya que las partes convendrían en ir sólo una ronda contra la otra a fin de ahorrar gastos de litigación de ambas; reside en que aproxima las ofertas de arreglo a lo que sería la evaluación judicial del caso si fuera litigado hasta la sentencia. Véase Bruce L. Hay, "Some Settlement Effects of Preclusion", 1993 *U. Ill. L. Rev.,* 21.

15. ¿Cree que es más o menos difícil arreglar un caso, y en términos más o menos ventajosos para el demandante, si el juicio se plantea contra varios demandados conjunta y separadamente responsables de cualquier cuasidelito cometido contra el demandante? Véase Kathryn E. Spier, "A Note on Joint and Several Liability: Insolvency, Settlement, and Incentives", 23 *J. Legal Stud.,* 559 (1994).

16. La Regla 11 de las Reglas Federales del Procedimiento Civil requiere que un abogado realice una indagación razonable sobre las bases factuales y legales de cualquier documento que él introduzca en el tribunal de instrucción. ¿Cuáles son las consecuencias probables de la regla para el costo y la cantidad de litigación?

17. Supongamos que algunos abogados son optimistas por naturaleza, mientras que otros son pesimistas. ¿A cuáles de ellos preferiría tener como su abogado en la fase previa a la instrucción de la litigación? ¿A cuáles de ellos preferiría tener en el juicio? En cuanto a su segunda respuesta, ¿sería diferente dependiendo de si el juicio se conduce bajo la regla estadunidense o la británica para determinar los honorarios de los abogados? Véase Oren Bar-Gill, "The Evolution and Persistence of Optimism in Litigation", 22 *J. Law, Econ. & Organization,* 490 (2005). ¿Existe entre los abogados una tendencia a mostrarse sobradamente seguros de sí? Véase Zev J. Eigen & Yair Listokin, "Do Lawyers Really Believe Their Own Hype, and Should They? A Natural Experiment", 41 *J. Legal Stud.,* 239 (2012).

XXIII. LA PRUEBA

XXIII.1. La prueba como búsqueda y como minimización del error

La ley de la prueba es el conjunto de reglas que determina cuál información, y cómo, podrá proveerse a un tribunal al que se pide que resuelva una disputa de hecho. El capítulo anterior destacó la importancia de la resolución correcta de tales disputas para un sistema de derecho económicamente eficiente, pero no discutió las reglas que estructuran el proceso de determinación de los hechos.

El proceso puede modelarse en dos formas en términos económicos; son equivalentes. La primera consiste en representarlo como un problema de búsqueda, análogo al de la búsqueda de un bien de consumo durable, donde la respuesta correcta a la cuestión (digamos) de si X le disparó a Y corresponde a la elección maximizadora de la utilidad entre dos marcas de un bien de consumo. El proceso de obtener, filtrar, reunir, presentar y (para quien establece los hechos) ponderar las pruebas confiere beneficios e incurre en costos. (Debe distinguirse entre los beneficios y los costos sociales y privados, pero eso vendrá más adelante.) Los beneficios son una función positiva de la probabilidad (p) de que, si la prueba es considerada por quien establece los hechos, el caso se decidirá correctamente, y de lo que esté en juego (S) en el caso. Supongamos, para simplificar, que los beneficios son simplemente el producto de los dos términos, en adelante pS, donde p es una función positiva de la cantidad de la prueba (x); entonces la expresión plena para los beneficios de la búsqueda es $p(x)S$. Con una prueba suficiente, p podría ser igual a 1, lo que significa que un juicio produciría seguramente el resultado correcto. Los costos del juicio (c) son también una función positiva de la cantidad de la evidencia.

Con estos supuestos, los beneficios netos ($B(x)$) de la búsqueda de pruebas en un caso están dados por $B(x) = p(x)S - c(x)$. La cantidad óptima de búsqueda —la cantidad que maximiza los beneficios netos— es entonces la cantidad que satisface $p_xS = c_x$, donde los subíndices denotan derivadas (véase antes VI.1, nota 2). En palabras, la búsqueda debiera llevarse hasta el punto en que el costo marginal se iguala al beneficio marginal. La cantidad de prueba en el punto óptimo será más alta entre mayor sea lo que está en juego en el caso, menor sea el costo de obtención de las pruebas, y mayor sea el efecto de las pruebas sobre el aumento de la probabilidad de un resultado correcto. Las condiciones que garantizan que este óptimo exista son que no haya economías de escala en la búsqueda de pruebas y que, en la medida que se ob-

tienen más y más pruebas, disminuya el efecto de las pruebas adicionales sobre el resultado del caso, condición que tenderá a obtenerse si el buscador inicia la búsqueda con la prueba más probatoria: el procedimiento racional, a menos que obtener esa prueba sea particularmente costoso.

La forma alternativa consiste en modelar la búsqueda de pruebas como un proceso de minimización del costo. Sea que p denote ahora la probabilidad de un resultado erróneo en lugar de uno correcto, y que pS denote el costo del error (la probabilidad del error ponderada por lo que está en juego). Supongamos que $p = 0.1$, lo que implica que, en promedio, uno de cada 10 casos se decidirá incorrectamente. Si lo que está en juego en esos casos son en promedio 100 000 dólares, el costo esperado del error es entonces de 10 000. (Es arbitrario el supuesto de que pS es igual a los costos sociales del error. Pero, como veremos más adelante, es una conjetura razonable que el costo social de un resultado erróneo aumentará generalmente con lo que esté en juego en el caso.) Debe emprenderse la búsqueda hasta el punto en que la última muestra de evidencia obtenida dé por fruto una reducción en los costos del error igual al costo de obtenerla.

En ocasiones ese costo puede ser bastante indirecto. Consideremos la Regla 407 de las Reglas Federales de Evidencia (FRE), según la cual es inadmisible la prueba de que, después del accidente que originó la demanda del demandante, el demandado reparó la condición que causó el accidente. Si se permitiera la introducción de tales pruebas en la instrucción de un caso, se desalentarían las reparaciones volviéndolas más costosas (siendo el costo adicional el aumento generado por tales pruebas en el costo esperado de la litigación para el demandado sobre el accidente), lo que aumentaría el riesgo de accidentes futuros y, por ende, los costos esperados del accidente. Contra esto se puede argumentar que el tomar una precaución una vez ocurrido un accidente que ésta habría prevenido justifica inferir que tener ese cuidado antes del accidente se habría justificado por los costos. Pero esto depende de las razones por las que se tomó la precaución subsecuente. Éstas podrían incluir (aparte de que ello siempre se justificaba por los costos) las presiones regulatorias, las preocupaciones por la publicidad adversa, un cambio reciente en el costo o la eficacia de la precaución, o también que el accidente revelara la existencia de un daño remediable que ninguna persona razonable hubiera sospechado. La exploración de estos asuntos desestabilizaría el juicio y reforzaría el problema del sesgo en retrospectiva. Las medidas de remedio subsecuentes son admisibles en tanto que muestran que había medidas preventivas de carácter factible que el demandado podría haber tomado (cuestión diferente a la de una adopción justificada por los costos), en el caso de que las medidas adoptadas tiempo después hubiesen existido antes del accidente; y, dada la excepción de viabilidad, la regla impide que el demandante argumente que adoptar las medidas equivale a admitir negligencia o defecto.

Lo que está en juego en un caso es cuando mucho una medida aproximada de los beneficios sociales derivados de la reunión de pruebas adicionales. Imaginemos una disputa sobre la responsabilidad según una ley que se derogó después del surgimiento de la disputa pero que, no habiendo sido derogada retroactivamente, todavía gobierna la disputa. Si está implicado mucho dinero, la inversión *privada* óptima en la reunión de pruebas podría ser muy grande porque la victoria conferirá o conservará rentas económicas en verdad sustanciales. Pero los beneficios sociales derivados de una decisión correcta podrían ser nulos. (O quizá no: la expectativa de que cualquier disputa surgida en los términos de la ley se resolviera por métodos correctos, pasase lo que pasase, podría haber inducido un comportamiento eficiente cuando la ley estaba vigente, y el respeto guardado a tal expectativa podría ser necesario para inducir el comportamiento eficiente en los términos de leyes actualmente vigentes.) Otro ejemplo sería el de la litigación por la propiedad de un tesoro enterrado; tal litigación generaría los mismos costos motivados por la renta que una carrera para encontrar el tesoro enterrado en un régimen de propiedad de quien encuentre el tesoro. Véase antes III.2. Así, si bien las partes podrían invertir deficientemente en la búsqueda de pruebas en algunos casos porque la corrección de la adjudicación confiere beneficios a quienes no son partes aumentando la eficacia disuasiva de la ley, pueden invertir excesivamente en otros casos por razones de búsqueda de renta. Véase antes XXII.19. Pero, en general, entre mayor sea lo que esté en juego en un caso, mayores serán los beneficios sociales y privados de una decisión correcta. La inexactitud que reduce la disuasión y, por lo tanto, el cumplimiento de la ley impondrá mayores costos sociales entre mayor sea el caso. Es más importante la desistencia de los derrames de petróleo de miles de millones de dólares que la de los de millones de dólares resultantes de la negligencia.

Segundo, las inversiones en pruebas podrían generar beneficios que van más allá de la alteración del fallo de casos particulares. Veamos el caso simple en que el fallo de un juicio depende sólo de la razón de las inversiones de cada parte. Entonces, una reducción proporcional de los gastos de cada parte no alterará el fallo. Pero bien podría reducir la cantidad de información generada para la consideración del tribunal, y aumentar la varianza del fallo efectivo por oposición al fallo esperado, además de la probabilidad de una apelación al disminuir la confianza en el apego del juicio a la verdad precisa.

XXIII.2. Comparación ulterior de los modelos acusatorio e inquisitivo

Los beneficios y los costos de la búsqueda de pruebas y, por lo tanto, la clase y la cantidad óptimas de tal búsqueda varían con el tipo del buscador; es la diferencia de quién busca las pruebas lo que distingue principalmente al sistema acusatorio del sistema inquisitivo (véase antes XX.1).

El sistema inquisitivo asigna al juez responsabilidades mucho más pesadas por la búsqueda de pruebas que el sistema acusatorio; y podría parecer que el buscador-juez sería un buscador de carácter extremadamente eficiente gracias a la selección, el adiestramiento y la experiencia. Pero quizá no sea así. Dado que es difícil evaluar la determinación de los hechos legales y, por ende, criticar a un juez por haber hecho determinaciones erróneas, o elogiarlo por hacerlas bien, será limitado el incentivo del juez para esforzarse a fin de hacer un buen trabajo. Además, la cantidad de la búsqueda realizada dependerá del número de jueces y del personal judicial auxiliar, y ese número podría determinarse sin tomar muy en cuenta la cantidad de búsqueda socialmente óptima.

En el proceso acusatorio ejemplificado por el moderno juicio civil por jurados de los Estados Unidos, la búsqueda de pruebas la realizan por separado los abogados de las partes, y las pruebas se presentan a un tribunal múltiple no experto, *ad hoc*, para la decisión. En virtud de que los abogados de instrucción son compensados directa o indirectamente sobre la base del éxito en un juicio, es muy grande su incentivo para encontrar pruebas favorables a su cliente y descubrir las fallas de las pruebas del contrario; y si se trata de un caso que implique grandes intereses monetarios en juego, recibirán de sus clientes amplios recursos para obtener y refutar las pruebas. Así, habrá una tendencia para, al menos, llevar la búsqueda de la evidencia hasta su punto óptimo.

Pero la cantidad de la búsqueda no es impulsada sólo por lo que está en juego, sino también por el efecto probable de la última prueba sobre el fallo. Esto implica que, en igualdad de las demás circunstancias, se obtendrán más pruebas entre más complicado sea el caso.[1] Si éste está cargado para uno de los lados, la prueba adicional, aunque en sí misma tenga gran valor probatorio, podría no tener ningún efecto sobre el fallo, de modo que no valdría la pena su presentación. Más aún, la prueba adicional inducida por lo cerrado del caso puede también no tener ningún efecto sobre el resultado. Supongamos que la parte A puede aumentar 1% la probabilidad de un fallo favorable añadiendo un poco más de prueba a un costo de x. Y supongamos que su contrario, B, puede anular ese desplazamiento de 1% en favor de A añadiendo otro elemento de prueba favorable a B, también a un costo de x. Si cada parte aporta su prueba adicional, se incurrirá en un costo de $2x$ sin lograr cambio alguno en el fallo esperado. Es cierto que, si A puede prever la respuesta de B, no tendrá ningún incentivo para aportar la prueba adicional. Y A y B se beneficiarán conjuntamente si se ponen de acuerdo para no aportar la prueba, y normalmente se respetarán los acuerdos de las partes para limitar las pruebas. Si los abogados fueran agentes perfectos de sus clientes, tales acuerdos

[1] Véase Avery Katz, "Judicial Decisionmaking and Litigation Expenditure", 8 *Int'l Rev. Law & Econ.,* 127 (1988).

serían más comunes de lo que son. Además, los jueces limitan la cantidad de pruebas que se permite aportar a las partes, como veremos más adelante.

El carácter competitivo del proceso de búsqueda y la presentación de los resultados a un cuerpo de jueces aficionados (los jurados) que no participa en la obtención de las pruebas imparten un carácter *retórico* al sistema acusatorio. La retórica, como explicó Aristóteles en el tratado sobre el tema, es el conjunto de técnicas para inducir creencias en cuestiones que implican la incertidumbre irremediable. Una dimensión importante de la retórica eficaz consiste en volver creíble al orador. Recuérdese que un bien es un bien de creencia si el consumidor no puede determinar fácilmente su calidad mediante la inspección o aun el uso (véase antes IV.7). Cuando no puede determinarse fácilmente el producto, la atención de un comprador se desplaza hacia los insumos de ese producto; así se explica el hincapié que se hace en la reputación del vendedor de un bien de creencia, o en la confiabilidad de un testigo cuyo testimonio no puede evaluarse directamente.

La importancia de la credibilidad y los incentivos de los abogados para aumentar la credibilidad de sus testigos explican el hincapié de los sistemas acusatorios en los interrogatorios y la refutación, y la sospecha correspondiente de las pruebas basadas en lo que se conoce de oídas, las que, definidas funcionalmente, son simplemente un testimonio que no está sujeto a un interrogatorio. El testigo que presenta en el tribunal una evidencia fundada en algún rumor puede ser sometido a un interrogatorio, pero no el declarante que nunca comparece ante la corte y cuyo "testimonio" está meramente repitiendo el testigo.

A menudo se malentiende la importancia del interrogatorio y, en consecuencia, se subestima su valor social, porque no se considera el efecto disuasivo del derecho a hacer sistemática, metódicamente pregunta tras pregunta. En virtud de que los interrogatorios *pueden* destruir la credibilidad de un testigo, raras veces lo hacen así en la práctica, de modo que son denigrados erróneamente. El testigo cuya credibilidad sería destruida por un interrogatorio no será llamado a declarar en absoluto, o tratará de eliminar el aguijón del inquiridor reconociendo en el examen directo los hechos que, según podría esperarse, éste estaría sacando a relucir con saña.

El sistema acusatorio parecería menos eficiente que el inquisitivo porque implica a dos o más buscadores —los abogados de las partes contrarias—, en lugar de uno solo (el juez). Y, en virtud de que los beneficios privados de la búsqueda de pruebas podrían ser mayores o menores que los beneficios sociales, la privatización de la búsqueda (como ocurre en el sistema acusatorio) podría generar demasiadas pruebas, o demasiado pocas desde un punto de vista social, como hemos visto, mientras que en principio —obviamente una salvedad enorme— el juez inquisitivo puede continuar su búsqueda de pruebas hasta que llegue al punto en que el costo marginal y el beneficio marginal se intersequen y pueda detenerse allí mismo. Con todo, el juez del sistema acu-

satorio puede limitar la cantidad de la búsqueda de los abogados, no sólo reduciendo el descubrimiento anterior a la instrucción, fijando una fecha temprana para el juicio y limitando la duración de la instrucción, sino también excluyendo ciertas pruebas del juicio bajo la autoridad de diversas reglas de la prueba. La función de las reglas de la prueba que limitan los costos externos generados por un sistema acusatorio es una de las razones por las que tales reglas son menos importantes en un sistema inquisitivo.[2]

Es más importante el hecho de que el carácter de competencia del proceso acusatorio otorga a los buscadores de pruebas (los abogados) un incentivo mayor para buscar con entusiasmo, y también un incentivo mayor para encontrar las fallas en las pruebas de la otra parte, que lo que ocurre con un sistema en que el juez es el principal o el único buscador.[3] En virtud de que la parte que tiene la posición objetivamente más fuerte podrá obtener pruebas favorables para ella a un costo menor que el necesario para que la parte contraria pueda obtener pruebas favorables,[4] el sistema competitivo de la reunión de pruebas tiende a favorecer a la parte que ganaría en un mundo libre de errores. También facilita obtener inferencias confiables a partir de las brechas en las pruebas. Si una de las partes debe ser capaz de obtener pruebas favorables para ella, si es que hay alguna, a un costo bajo, el hecho de que no presente tales pruebas permitirá que quien determina los hechos infiera que la parte no tiene tales pruebas y, por lo tanto, debe perder. El silencio se convierte en una señal.

Con una parte tan grande de la función judicial privatizada, el número de los jueces profesionales requeridos para equipar a los tribunales que emplean el sistema acusatorio es mucho menor que en los sistemas inquisitivos.[5] Alternativamente, la decisión de tener un número pequeño de jueces eleva el costo de su búsqueda, de modo que la función de búsqueda recae en otros, los abogados y los jurados, quienes podrían considerarse buscadores excesivamente costosos (en relación con los beneficios) si los jueces fuesen abundantes y la búsqueda judicial fuese, por ende, barata.

[2] Véase Gordon Tullock, *Trials on Trial: The Pure Theory of Legal Procedure,* 151-157 (1980).

[3] Véase Mathias Dewatripont y Jean Tirole, "Advocates", 107 *J. Pol. Econ.,* 1 (1999).

[4] Idealmente, el costo de la presentación de pruebas favorables debiera ser infinito para la parte que merece perder. Si así fuera, el hecho de que las partes tengan incentivos para mentir en un régimen de reunión competitiva de las pruebas no conduciría a resultados erróneos. Véase Chris William Sanchirico, "Games, Information, and Evidence Production: With Application to English Legal History", 2 *Am. Law & Econ. Rev.,* 342 (2000).

[5] Por ejemplo, la razón de abogados a jueces es de 54.59 a 1 en los Estados Unidos, mientras que en Francia es de 6.07 a 1, en Alemania de 6.86 a 1 y en Suiza de 2.86 a 1. Richard A. Posner, *Law and Legal Theory in England and America,* 28 (1996) (cuadro 1.1).

XXIII.3. UNA REVISIÓN DEL JUICIO POR JURADOS

Para complementar nuestra discusión anterior de los jurados (véase antes XXII.16), consideremos algunas características de los juicios por jurados desde el punto de vista de la teoría económica de la prueba.

Un juez puede estar endurecido por los casos, de modo que sea menos probable que atienda a los detalles de un caso nuevo. Supongamos que, por haber presidido en muchos casos similares, el juez avizora al inicio de un caso nuevo que la probabilidad de que el acusado sea culpable es de 100 a 1. Este juez tendrá escaso incentivo para prestar mucha atención a las pruebas presentadas en el juicio, porque las pruebas de la culpa del acusado no modificarán su juicio, mientras que las pruebas de la inocencia del acusado, a menos que sean extremadamente poderosas, no cambiarán los prejuicios del juez lo suficiente para inducirlo a absolver al acusado. El juez está siendo racional, pero cuando el patrón se vuelve conocido los litigantes ya no tendrán ningún incentivo para presentar muchas pruebas. (Para entender esto, piénsese en el caso extremo en el que el juez se ha convencido irrevocablemente del fallo correcto del caso antes de que se presente prueba alguna.) Finalmente, la corrección del proceso de litigación se verá gravemente comprometida, a medida que los prejuicios del juez, formados sobre la base de juicios en los que las partes presentaron muchas pruebas, se vuelven menos y menos correctos. El problema se agravará si la gente está afectada por el "sesgo de la confirmación", lo que quiere decir que tiende a interpretar las pruebas en la forma más congruente con sus prejuicios.[6] El hecho de que el juez vitalicio que gana un sueldo fijo no pague ningún castigo por sucumbir a este sesgo mina su resistencia ante él. Los peligros derivados de los prejuicios fuertes son menos agudos en el juicio por jurados.

Dado su carácter acusatorio y la necesidad de presentar todas las pruebas al mismo tiempo (un jurado no puede estar reunido indefinidamente, mientras que un juez puede tratar un caso en etapas a lo largo de un periodo indefinido y demorar la expedición de su decisión hasta mucho tiempo después de que el juicio haya terminado), el público puede monitorear un juicio por jurados estadunidenses a un costo de información menor del que se requeriría para monitorear un proceso inquisitivo modelado a la manera de una investigación policiaca. Este monitoreo es particularmente importante en el ambiente de los Estados Unidos porque la mayoría de los jueces estadunidenses son elegidos y no designados y sus sueldos judiciales están muy por debajo de los costos de oportunidad de los abogados más hábiles, lo que genera dudas acerca de la calidad judicial por ambas razones.[7] Las preocupaciones

[6] Véase Matthew Rabin, "Psychology and Economics", 36 *J. Econ. Lit*., 11, 26-28 (1998).

[7] Aunque obtener una inferencia adversa a partir del diferencial salarial se complica por dos factores. Primero, si los sueldos judiciales fuesen mayores, el resultado podría ser un incentivo

por la calidad estrechan a su vez la brecha de la competencia percibida entre los jueces y los jurados, y así reducen los costos esperados del error de los juicios por jurados en relación con los juicios sin jurado. Esa brecha se estrecha más aún por el hecho de que, cuando los jueces son elegidos, los litigantes frecuentes, como las compañías de seguros, y los abogados especializados en la instrucción, como los de la barra de demandantes de cuasidelitos, tienen fuertes incentivos para canalizar sus aportaciones de campaña hacia los jueces que favorecen a sus intereses. Estos incentivos menguan, y el efecto de las contribuciones sobre las decisiones disminuye, si la responsabilidad de la decisión se comparte con jurados.

Está claro que un juez tiene una ventaja comparativa sobre un jurado cuando se trata de decidir cuestiones de derecho. Pero no siempre es clara la línea divisoria entre una cuestión de hecho y una cuestión de derecho. La negligencia es un concepto jurídico, ¿pero si un demandado fue negligente es una cuestión jurídica o de hecho? La respuesta habitual es que es ambas cosas a la vez o ninguna de ellas, y quizás esto se describa mejor como una cuestión "mixta" de derecho y de hecho, o una cuestión acerca de la aplicación de un concepto jurídico a un conjunto de hechos. El análisis económico sugiere, empero, que es una pura cuestión de hecho, de modo que la ley tiene razón cuando la deja en manos del jurado sujeta a la misma revisión deferente que otras determinaciones fácticas realizadas por los jurados. Esto se entiende muy fácilmente cuando se expresa el concepto de la negligencia en los términos de la Fórmula de Hand (véase el capítulo VI). Toda determinación requerida para aplicar la fórmula tiene un carácter de hecho antes que legal: estimar *B, P* y *L,* multiplicar *P* y *L,* y determinar si *PL* es mayor que *B.* No se requiere ningún conocimiento legal para hacer ninguna de estas determinaciones (como ocurriría si se pidiera al jurado que decidiera qué es la negligencia), que cuando se hacen responden a la interrogante de si el demandado fue negligente.

Aunque gran parte de la oposición al juicio civil por jurados se basa en anécdotas, los oponentes pueden señalar algunas pruebas empíricas genuinas, en particular la prueba de que la tasa de condenas es menor en los juicios sin jurado que en los juicios por jurados. En la mayoría de los estados de la Unión Americana, la decisión que se tome en un caso penal para encomendar la instrucción a un juez o a un jurado corresponde por entero al demandado. Si los jurados son determinadores de la culpa menos acertados que los jueces, los demandados inocentes escogerán ser juzgados por jueces antes que

mayor por parte de los políticos a considerar los puestos de jueces como un empleo óptimo (lo mejor de lo mejor) a alcanzar merced al patronazgo; el puesto de juez valdría más para el amigo o partidario del político y, por lo tanto, más para el político. Segundo, los sueldos judiciales, como los militares, están deprimidos por el monopsonio. Sólo hay un empleador de jueces federales en los Estados Unidos, así como sólo hay un empleador de personal militar. Por lo tanto, si se quiere ser un juez federal, o un soldado, no se puede confiar en que la competencia entre empleadores potenciales otorgue un sueldo comparable al que se podría obtener en un empleo alternativo.

correr el riesgo de un error del jurado, mientras que los demandados culpables escogerán ser juzgados por jurados, esperando *que haya* un error; y entonces la tasa de absoluciones será mayor en los juicios sin jurado, como en realidad ocurre.[8]

XXIII.4. La carga de la prueba

La carga de la prueba tiene dos aspectos. El primero de ellos sólo es importante en un sistema acusatorio, donde el tribunal no participa en la búsqueda de pruebas. Ésta es la carga de la producción (presentación) de pruebas para el tribunal, por oposición a la carga de persuadir al tribunal de que debemos ganar el caso. Si no se cumple con cualquiera de las dos cargas, ello significa que la parte a quien corresponde esa carga pierde. Las dos cargas están interconectadas; por ejemplo, la carga de la persuasión determina generalmente quién tiene la carga de la producción. La carga del demandante en un caso civil ordinario consiste en demostrar que su posición es más probablemente correcta que incorrecta. En otras palabras, si al final del juicio piensa el jurado que el demandado debe ganar o no sabe cuál de las partes debe ganar —las pruebas parecen estar niveladas—, el demandante pierde. Esto vuelve muy probable que pierda el demandante que no ofrece ninguna prueba; por lo tanto, tiene sentido, como un medio de economización del tiempo del tribunal (y de disminución de la litigación ociosa), requerir al demandante, como una condición para que se inicie el juicio, que presente pruebas que, si se creyeran, harían que el demandante triunfara con el jurado, antes de que el demandado deba presentar prueba alguna.[9] Esto supone que el costo para el demandante de obtener estas pruebas no es desproporcionadamente mayor que el costo para el demandado de obtener pruebas en contrario (si tiene alguna). Pero el supuesto es razonable; los procedimientos modernos previos al juicio para el descubrimiento de pruebas en posesión de la parte contraria vuelven a los costos de la búsqueda de pruebas bastante simétricos.

¿Pero por qué, en el caso de empate, debiera corresponderle la victoria al demandado en vez de al demandante? Una razón de ello es que un juicio contra un demandado —por haberlo pescado cometiendo un delito— le causa más daño a su reputación que el que le causa a la reputación del demandante el intento malogrado de ganarse un juicio. Otra es que un juicio a favor del demandante impone un costo que se evita cuando el error se atribuye a la otra parte: el costo de reunir los elementos para el fallo.

La carga de la persuasión corresponde al demandante por lo que toca a la reclamación principal, pero recae sobre el demandado en lo tocante a las

[8] Hugo M. Mialon y Paul H. Rubin, "The Economics of the Bill of Rights", 10 *Am. Law & Econ. Rev.,* 1, 37 (2008); Gerald D. Gay *et al.,* "Noisy Juries and the Choice of Trial Mode in a Sequential Signalling Game: Theory and Evidence", 20 RAND *J. Econ.,* 196 (1989).

[9] Bruce L. Hay, "Allocating the Burden of Proof", 72 *Indiana L. J.,* 651 (1997).

defensas afirmativas tales como el consentimiento, la ley de limitaciones, la incapacidad y la *res judicata*; y las cargas de la producción se asignan de modo correspondiente. Sería particularmente ineficiente que se exigiera al demandante anticipar y presentar pruebas en contra del número indefinido de defensas que un demandado podría esgrimir en un caso dado. El demandante tendría que identificar y atacar defensas de cuya existencia quizá no tuviera conocimiento el demandado, así como algunas defensas que el demandado podría tener buenas razones tácticas o probatorias para no esgrimir o que simplemente optara por no esgrimir porque tenía una defensa claramente definitiva, de modo que no tenía que perder tiempo con otras. Por ejemplo, si la ley de limitaciones es una defensa plausible sólo en 5% de los casos, hacer que el demandado alegue y pruebe que su demanda se formuló en tiempo impondría costos sin beneficios correspondientes en 95% de los casos.

La carga de la persuasión es mucho más onerosa (la prueba más allá de una duda razonable contra la prueba por una mera preponderancia de la convicción) en un caso penal que en un caso civil. Esto tiene sentido económico. En el juicio civil típico no hay ninguna base para suponer que los errores de tipo I (falsos positivos, tales como la condena de una persona inocente o, en el contexto civil, declarar responsable al demandado erróneamente) imponen en promedio costos mayores que los errores de tipo II (falsos negativos, tales como la absolución errónea o la negación errónea de una reclamación meritoria).[10] Por lo tanto, en el caso civil habitual basta para justificar un veredicto en favor del demandante que la probabilidad de que su reclamación sea meritoria supere a la probabilidad de que no lo sea, por cualquier margen.

Los errores de tipo I son más graves que los errores de tipo II en los casos penales (sujeto ello a la importante calificación asentada antes en VII.12), de modo que se ponderan más fuertemente por la imposición de una carga pesada de persuasión impuesta al fiscal. En el capítulo VII vimos que, aun cuando la única sanción penal sea una multa, el costo pleno para el acusado condenado no aparecerá en ninguna otra parte de los libros sociales como un beneficio. Y, cuando no hay una multa, el único beneficio del castigo penal es su efecto disuasivo. Es improbable que sea positivo el efecto disuasivo del castigo a una persona inocente, y *podría* ser negativo, aunque no se puede decir nada más categórico. El costo esperado del castigo (*CE*) para una persona castigada es efectivamente la diferencia existente entre el costo esperado del castigo si comete un delito ($CE_g = p_gS$, donde p_g es la probabilidad del castigo si el

[10] El trueque de compensaciones y sacrificios que se da entre los errores de tipo I y II es una característica omnipresente del derecho de las pruebas. Consideremos las disputas acerca de si una rueda de presos para la identificación de la policía y los testigos es excesivamente "sugerente". Si las otras personas de la rueda se asemejan mucho al acusado, se minimiza la probabilidad de un error de tipo I (identificar erróneamente al acusado como el delincuente), porque el acusado no "se destaca". Pero aumenta la probabilidad de un error de tipo II (incapacidad para identificar al acusado como el delincuente que sí es), porque es más fácil confundir al acusado con las otras personas en la rueda.

acusado es culpable y *S* es la sentencia) y el costo esperado del castigo si el acusado no comete un delito ($CE_i = p_iS$, donde p_i es la probabilidad del castigo si el acusado es inocente, y *S* es la sentencia). Por lo tanto, $CE = p_gS - p_iS$; equivalentemente, $CE = (p_g - p_i)S$, lo que vuelve transparente que, si se impone el castigo al azar, de modo que la probabilidad del castigo sea la misma independientemente de la culpa (es decir, si $p_g = p_i$), el costo esperado del castigo por cometer el delito será cero. Entre más precisión se observe en el proceso para determinar la culpa, menos aleatorio será el castigo y, por lo tanto, será mayor el efecto disuasivo de la ley. Dicho de otro modo, la mayor precisión observada en la determinación de la culpa aumenta los rendimientos del hecho de ser inocente.

Pero el castigo que es verdaderamente aleatorio debe distinguirse del castigo que tiene sólo un componente aleatorio. Supongamos que es muy probable que una persona con antecedentes penales sea condenada por cualquier delito que se le impute en el futuro, aunque no lo haya cometido. Esto reducirá el efecto del derecho penal para disuadirla de cometer nuevos delitos, pero al mismo tiempo volverá más probable que se vea disuadida de cometer el primer delito, al aumentar su castigo esperado (de largo plazo). También hará que las personas con antecedentes penales se alejen de las actividades en las que podrían ser arrestadas y falsamente acusadas. Así pues, la condena de una persona inocente puede aumentar tanto como reducir la disuasión;[11] aun así, es probable que predomine el último efecto en la mayoría de los casos, pues si los encargados de cumplir la ley concentran sus limitados recursos en las personas más fáciles de condenar, independientemente de la culpa, el castigo esperado de quienes delinquen por primera vez disminuirá porque las autoridades estarán destinando menores recursos a su procesamiento.

Los errores del tipo I en los casos penales imponen costos adicionales porque el costo del encarcelamiento es sustancial. Por supuesto, ese costo se evita cuando una persona culpable es absuelta, aunque tal absolución reducirá la disuasión al reducir la probabilidad del castigo por el delito. Pero el efecto asimétrico del costo del encarcelamiento sobre las condenas y las absoluciones significa que probablemente se necesiten varias absoluciones erróneas para imponer un costo social igual al de una condena errónea. Ésa es una justificación económica para requerir una prueba más allá de toda duda razonable en un caso penal por oposición a un caso civil, y otra es la ventaja inherente que disfruta el fiscal en un caso penal, por comparación con un demandante privado en un caso civil. El gobierno tiene enormes recursos para la persecución. Puede repartir estos recursos entre los casos como lo desee, obteniendo declaraciones de culpabilidad mediante la amenaza de concentrar sus recursos contra cualquier acusado que se niegue a confesar, y utili-

[11] Véanse otros ejemplos de cómo la incorrección de la adjudicación puede aumentar paradójicamente el bienestar social en Michael L. Davis, "The Value of Truth and the Optimal Standard of Proof in Legal Disputes", 10 *J. Law, Econ. & Organization,* 343 (1994).

zando los recursos así conservados para apabullar al acusado ocasional que sí invoque su derecho a un juicio.[12] Esto es como la situación discutida en el capítulo X, en la que el acceso desigual a los mercados de capital puede hacer de los precios depredadores una estrategia racional. La analogía es mayor, y la estrategia probablemente más eficaz, cuando, como es el caso más común, el acusado no tiene recursos para contratar a un abogado, sino que depende de los abogados designados por el tribunal, quienes tienen escasa capacidad financiera excepto en los casos de pena capital. Aun el raro acusado penal al que le alcanza para contratar abogados será normalmente incapaz de igualar los recursos con los que el gobierno puede amenazar creíblemente en un caso. La carga de la prueba de la culpa más allá de una duda razonable es, pues, una compensación parcial (como proporcionar un abogado a los acusados indigentes) de la desigualdad de los recursos de las partes para reunir y presentar pruebas. En un sistema inquisitivo, donde la búsqueda es realizada por un juez presumiblemente desinteresado, se atenúa la necesidad de una carga de la prueba que sea mayor en un sistema penal que en un sistema civil.

Un factor de complicación es que los fiscales podrían ser también desinteresados en cualquiera de los sistemas, porque, al revés de lo que ocurre con los abogados privados, sus ingresos no están directamente ligados al éxito en la litigación. Pero la teoría económica, al igual que el sentido común y la observación, sugiere que el deseo de ganar, ponderado por lo que esté en juego en el caso (dicho aproximadamente, la sentencia si el acusado es condenado), es el argumento más importante de la función de utilidad de los fiscales, y que, de este modo, los fiscales tienen incentivos similares a los de los abogados privados. El empleo de fiscal es raras veces un empleo terminal; es un trampolín. Los empleadores futuros evaluarán a un fiscal por su éxito en la litigación, el que se verá como una función de su tasa de triunfos ponderada por la oposición que tuvo que superar para ganar; la oposición será de ordinario mayor entre más grave sea la ofensa imputada.

Cuando se les pide que expresen una prueba más allá de una duda razonable como una probabilidad de la culpa, los jueces seleccionan generalmente un número entre 0.75 y 0.95 (dependiendo del juez); las cuantificaciones del jurado son similares. Estas cifras podrían parecer perturbadoramente bajas, ya que implican que hasta una cuarta parte de los individuos condenados por un delito son inocentes. No es así. Entre mayor sea la tasa delictiva en relación con los recursos de los fiscales, más a fondo seleccionarán los fiscales los casos más fáciles de ganar, y éstos tenderán a ser extraídos de la cola de la distribución de los sospechosos que contiene a los que tienen mayores probabilidades de ser efectivamente culpables.[13] La pesada carga de la persuasión

[12] La elección de casos por las dependencias encargadas del cumplimiento de la ley se discute con mayor extensión después, en XXIV.3.

[13] Véanse algunas pruebas dramáticas de esto en J. Mark Ramseyer y Eric B. Rasmusen, "Why Is the Japanese Conviction Rate so High?", 30 *J. Legal Stud.,* 53 (2001).

y las otras ventajas procedimentales de los acusados penales aumentan el incentivo de los fiscales para ir tras los más culpables al volver difícil la condena de un acusado, a pesar de la disparidad de los recursos entre las autoridades del ministerio público y todos los acusados con excepción de los más ricos, a menos que el caso esté cargado en su contra. Si debido a la selección realizada por los fiscales sólo son inocentes 1% de las personas procesadas, entonces, aunque *todos* los acusados sean condenados, sólo 1% de las personas condenadas serán inocentes. Y ésa es una exageración. No todas las personas que son procesadas resultan condenadas, y normalmente es mucho más fácil para un acusado inocente que para un culpable crear suficientes dudas de la culpa en la determinación de los hechos para inducir a una absolución.

La selección estricta implica no sólo que algunos individuos culpables no son procesados, sino también que la mayoría de los individuos que son procesados y *absueltos* son efectivamente culpables. En el ejemplo anterior, si se supone una tasa de absolución de 10%, entonces 99% de los acusados que son absueltos habrían sido efectivamente culpables si la probabilidad de la absolución fuese aleatoria respecto de la inocencia y 90% si todos los acusados inocentes fuesen absueltos.[14] Esto implica que cuando las tasas delictivas aumentan más de prisa que los recursos del ministerio público, lo que genera un filtrado más fino aún en la selección de los casos que habrán de perseguirse, los tribunales o las legislaturas tendrían que disminuir las ventajas procedimentales de los acusados si desearan mantener el mismo equilibrio entre las probabilidades de condenar al inocente y de absolver al culpable. Esta observación sugiere una posible base no ideológica para el desplazamiento de la Corte Suprema en contra de los derechos de los acusados penales en los decenios de 1970 y 1980. Si tales derechos se hubieran preservado intactos, la elevación de las tasas delictivas en esa época (que superó ampliamente el aumento del número de las persecuciones[15]) habría facilitado que los acusados evitaran el castigo. Ello habría reducido a su vez el costo esperado del castigo y, por ende, aumentado más aún las tasas delictivas, a menos que hubiera un aumento compensatorio en la severidad del castigo para los (pocos) delincuentes que fueran aprehendidos y condenados.

[14] Entonces, de 10 000 acusados se presume que 100 (1%) son inocentes. Si 1 000 de los acusados son absueltos (10%) y la probabilidad de la absolución es la misma independientemente de la inocencia, en ese grupo de 1 000 sólo 1% serán inocentes y 99% culpables. Si se absuelve a los 100 inocentes, 900 de los acusados absueltos deberán ser culpables (90 por ciento).

[15] Entre 1960 y 1996, el "índice de la delincuencia" compilado por el FBI y reportado en la publicación anual del FBI, *Uniform Crime Reports*, creció casi siete veces. Durante el mismo periodo, los procesos penales crecieron sólo un tercio (véanse los informes anuales del director de la Oficina Administrativa de los Tribunales de los Estados Unidos). (Desde entonces, empero, el número de procesos penales a nivel federal se ha acrecentado agudamente y las tasas de delitos han bajado.) Para los procesos penales estatales sólo hay datos a partir de 1977. Entre esa fecha y 1994, el número de procesos disminuyó en realidad ligeramente (según la estimación del Proyecto de Estadísticas Judiciales del Centro Nacional para los Tribunales Estatales), en un periodo en el que el índice de criminalidad estaba creciendo un tercio.

Aunque la carga de la persuasión es mucho menor en un caso civil y la mayoría de los demandantes no operan con una restricción de recursos (gracias a los honorarios contingentes), no hay ninguna razón para suponer que se decida incorrectamente una fracción mayor de los casos civiles que de los penales. La carga de la persuasión tiene menos que ver con el número de los errores que con su distribución entre las partes. Ganan más demandantes civiles sin méritos que fiscales penales sin méritos, pero también ganan menos demandados civiles sin méritos que acusados penales sin méritos. Lo que vuelve probable que la mayoría de los casos, civiles o penales, se resuelvan en forma correcta es simplemente que de ordinario resulta más barato obtener pruebas persuasivas del lado de la verdad.

Ahora se acepta que, dado que todas las pruebas son probabilísticas no debieran excluirse sólo porque su apego a la verdad de los hechos pueda expresarse en términos probabilísticos que revelen una alta (aunque no de 100%) probabilidad de apego a los hechos, como en el caso de las huellas digitales y el ADN. Pero los tribunales se resisten a dar el paso siguiente y sostienen que, dada la moderada carga de la persuasión en los casos civiles, la probabilidad explícita de que sea cierta la prueba esencial del demandante sólo necesita exceder de 50%. Supongamos que el demandante es golpeado por un autobús, y que se sabe que 51% de los autobuses que circulan por el camino donde el demandante fue golpeado son propiedad de la compañía de autobuses A y 49% son propiedad de la compañía B. La víctima demanda a A y pide que se falle sobre la base de esta estadística exclusivamente; no presenta ninguna otra prueba. Si el demandado tampoco presenta prueba alguna, ¿debiera permitirse que un jurado falle en favor del demandante? La respuesta de la ley es "no" y se puede justificar por el escepticismo ante el hecho de que la estadística referente a la propiedad de los autobuses sea la única prueba que el demandante puede obtener. Si la estadística es la única prueba del demandante, la inferencia que debe obtenerse no es que haya una probabilidad de 51% de que fue un autobús propiedad de A el que atropelló al demandante, sino que el demandante investigó y descubrió que fue efectivamente un autobús propiedad de B (y digamos que B está a prueba de los fallos, de modo que no vale la pena demandarlo) o de que simplemente no se ha molestado en realizar una investigación. Si la primera alternativa es cierta, el demandante deberá perder, por supuesto; y, dado que eso puede ser cierto, la probabilidad de que el demandante haya sido atropellado por un autobús propiedad de A es menor a 51%. El demandante debiera perder aun si fuese cierta la segunda alternativa (que no se molestó en investigar). No debiera exigirse al tribunal que gaste sus escasos recursos de tiempo y esfuerzo en un caso, hasta que el demandante haya realizado una búsqueda suficiente para indicar que es razonablemente probable que un gasto de recursos públicos rinda un beneficio social significativo. Eso está implícito en la decisión discutida antes de imponer la carga de presentación de la prueba al demandante y no al demandado.

Supongamos que ambas partes realizan una investigación a fondo, pero no pueden obtener ninguna prueba adicional acerca de la propiedad del autobús. Ya no hay ninguna base para sospechar que el demandante crea realmente que lo atropelló un autobús propiedad de la compañía B, o para castigarlo por no haber investigado más. Pero todavía hay algunas razones para la desestimación. Es probable que los costos de la instrucción de esos casos superen a los beneficios sociales. Si se instruyen 1 000 de tales casos, podemos esperar 510 decisiones correctas (es decir, 510 decisiones en las que el demandado fue efectivamente el victimario) y 490 decisiones incorrectas, mientras que puede esperarse que desestimar todos los casos produzca 490 decisiones correctas y 510 erróneas. Los beneficios sociales de las 20 decisiones adicionales correctas que produciría la instrucción de los 1 000 casos —beneficios en la desistencia más perfecta de los accidentes por negligencia— serían probablemente menores al costo social de 1 000 juicios. Peor aún: los beneficios sociales pueden ser negativos. Si la ley establece que la porción del mercado de la compañía A es suficiente para la responsabilidad y no se dispone de ninguna otra prueba, entonces se derrumbará el incentivo de la compañía B para tener cuidado porque sabrá que A será responsable de cualesquiera accidentes que cause B.[16] B habrá logrado externalizar los costos de sus accidentes.

XXIII.5. El error inocuo

La Regla 103(a) de las Reglas Federales de Evidencia establece entre otras cosas que un fallo que admita o excluya una prueba no puede servir de base a la moción para entablar un nuevo juicio o para revocar una apelación, "a menos que se vea afectado un derecho sustancial de la parte [contra la que se dictó el fallo]". En otras palabras, los errores inocuos serán descartados. La doctrina del error inocuo no se limita a los fallos sobre las pruebas pero, debido en parte a que tales fallos son tan frecuentes en un juicio, son los candidatos más frecuentes al tratamiento de esta doctrina. Además, los fallos sobre la inocuidad dependen de supuestos acerca del entendimiento de la persona que determina los hechos.

El error inocuo desempeña un papel particularmente importante en las apelaciones penales. En virtud de que la mayoría de los demandados penales no le pagan a su abogado, apelarán su condena o su sentencia aunque sea escasa la probabilidad de la revocación; si bien los beneficios esperados de la apelación podrían ser escasos, los costos esperados de la misma son de cero. Por lo tanto, los errores menores figuran prominentemente en las apelaciones penales, y esto pone en claro la necesidad de una doctrina del error ino-

[16] Véase Eric B. Rasmusen, "Predictable and Unpredictable Error in Tort Awards: The Effect of Plaintiff Self-Selection and Signaling", 15 *Int'l Rev. Law & Econ.,* 323 (1995).

cuo para eliminar las prisiones preventivas que serían puros costos y nada de beneficios. Pero, a pesar del atractivo de sentido común (y económico) que posee el principio del error inocuo, a menudo podría conferir una ventaja inmerecida, o por lo menos no buscada, a los fiscales.[17] La regla está sesgada en su favor porque el tribunal de apelación, no habiendo presenciado la instrucción (y especialmente no habiendo observado al jurado), carece de buena información para evaluar la probabilidad de que los errores afectaran al fallo. El tribunal de apelación, sabiendo poco o nada sobre la composición real del jurado en el caso, basará forzosamente su evaluación del carácter inocuo en el probable efecto que tengan los errores sobre el jurado *promedio*, mientras que el fiscal podría saber que el jurado particular ante el que procesó al acusado tiene una propensión a absolver por encima de la media, de modo que había tenido que usar medios impropios para obtener una condena.

Si el fiscal induce al juez de instrucción a cometer un error en su favor, hay tres consecuencias posibles: una condena seguida de la confirmación; una condena seguida de la revocación y luego un nuevo juicio, y la absolución, que no genera ninguna ganancia para el fiscal y lo deja con una pérdida neta. La ganancia neta esperada por el fiscal del error es entonces la suma de su ganancia si hay una condena confirmada, su ganancia (menos sus costos adicionales) si hay una condena revocada seguida de un nuevo juicio que termina en una sentencia confirmada, y su pérdida (más sus costos adicionales) si el nuevo juicio termina en una absolución, con la primera ganancia descontada por la probabilidad de la confirmación y la segunda por la probabilidad de la revocación.

El error en favor del fiscal aumenta su ganancia neta esperada al aumentar la probabilidad de una condena, pero la disminuye al aumentar la probabilidad de una revocación de la condena y al aumentar los costos del fiscal. La posibilidad de que el efecto neto sea positivo no puede ser excluida y aumenta si consideramos la sustituibilidad de la violación de las reglas procedimentales o probatorias en lugar de otros insumos de la obtención de una condena. Porque entonces el efecto del uso de tácticas abusivas, además de incluir como antes la reducción de la probabilidad de una condena que permanecerá y, por ende, aumentar así la probabilidad de tener que incurrir en el costo de un segundo juicio, disminuye también el costo del primero.

XXIII.6. La prueba del balanceo de la Regla 403

Las Reglas Federales de Evidencia definen la relevancia como "el hecho de tener cualquier tendencia a hacer que la existencia de cualquier hecho que es importante para determinar la acción sea más (o menos) probable de lo que

[17] Como una cuestión práctica, sólo los fiscales se benefician de la regla en casos penales: tan raras son las apelaciones de los fiscales.

habría sido sin la prueba" (Regla 401); pero aun la prueba relevante podría ser excluida del juicio "si su valor probatorio se ve sustancialmente superado por el peligro del prejuicio injusto, la confusión de los problemas o el engaño del jurado, o por consideraciones de una demora indebida, desperdicio de tiempo o presentación innecesaria de pruebas acumuladas" (Regla 403).

Al requerir una comparación explícita del beneficio y el costo, la Regla 403 es fundamental para un análisis económico del derecho de la prueba en una forma muy semejante a como la Fórmula de Hand es fundamental para el análisis económico del derecho de los cuasidelitos. Esta regla establece una fórmula de costo-beneficio para decidir la cuestión más común del derecho de la prueba, que es la decisión de admitir o desechar una prueba. También puede usarse para evaluar reglas de la prueba particulares, así como la Fórmula de Hand se usa en el análisis económico del derecho como un estándar para evaluar reglas específicas del derecho de los cuasidelitos. Con todo, a diferencia de la Fórmula de Hand, la Regla 403 hace hincapié en la escala ("*sustancialmente* superado"); la razón estriba en que es de suponer que quien determina los hechos, juez o jurado, tenga alguna habilidad para hacer caso omiso de los rasgos perjudiciales que tenga la prueba provista de valor probatorio.

La Regla 403 utiliza tres razones diferentes para excluir las pruebas relevantes: *1)* la emocionalidad (una fuente del "prejuicio injusto" y del "engaño al jurado"), *2)* la excesiva carga cognitiva (la "confusión" y otras formas del "engaño al jurado") y *3)* la "pérdida de tiempo" (que parece sinónimo de la "demora indebida" y la "presentación innecesaria de pruebas acumuladas"). Las dos primeras razones se relacionan con las limitaciones cognitivas de quien determina los hechos, de modo que van del lado de los beneficios de las pruebas en la determinación de la verdad, mientras que la tercera va del lado de los costos. Deben distinguirse dos tipos distintos de la limitación cognitiva. El primero, a menudo llamado "racionalidad restringida", surge del hecho de que los individuos no tienen costos nulos en la absorción y el análisis de la información, de modo que enfrentan problemas de sobrecarga. Este tipo de la limitación cognitiva (razón 2) es enteramente congruente con la racionalidad, que no presupone costos nulos de la adquisición y el procesamiento de la información. El segundo tipo de la limitación cognitiva es el resultado de las distracciones emocionales (razón 1). La ocultación de pruebas al jurado es una alternativa para lo que podrían ser esfuerzos que consumen mucho tiempo y son ineficaces a fin de aumentar y dejar libres de sesgos las capacidades cognitivas del jurado. Los jurados no tienen incentivos monetarios para hacer un trabajo cuidadoso; al ocultarles materiales probatorios que volverían su trabajo más difícil aún, y obligarlos así a realizar mayores esfuerzos mentales sin compensación alguna, las reglas de la prueba reducen los costos en que incurren los jurados para llegar a un fallo correcto.

La razón *2)* interacciona también con *3)* (la pérdida de tiempo): la repetición y la prolongación pueden hacer más difícil que el encargado de deter-

minar los hechos llegue a un fallo correcto, además de aumentar los costos directos del juicio. A medida que se introducen más y más pruebas, las pruebas adicionales, aunque sean relevantes, tienden a ser un desperdicio en el sentido de que rinden beneficios decrecientes por lo que toca a la mayor precisión sin una disminución correspondiente de los costos y a resultar confusas en el sentido de que en efecto disminuyen la precisión. Esta observación sugiere que la extensión óptima de un juicio por jurados debe ser muy breve en la mayoría de los casos; los beneficios de las pruebas adicionales tienden a declinar aceleradamente, mientras que los costos son constantes, o aun crecientes, en la medida que los litigantes se extravían más y más.

XXIII.7. La prueba del carácter

La Regla 404 de las Reglas Federales de Evidencia excluye (con diversas excepciones) las pruebas del carácter de una persona cuando se utilizan para demostrar que probablemente actuó "de acuerdo con su carácter" en la ocasión implicada en la litigación. La consecuencia principal es la exclusión de las pruebas de que un acusado penal tiene antecedentes penales a menos que testifique, como veremos más adelante. Tales pruebas son relevantes porque una persona que ha cometido un delito en el pasado ha indicado así una propensión a cumplir por debajo del promedio con las leyes penales. Pero esto tiene un valor probatorio débil, sobre todo porque los delincuentes reincidentes son castigados más severamente que quienes delinquen por primera vez, en parte precisamente para contrarrestar cualquier propensión mayor a la comisión de delitos que su delito o delitos anteriores hayan revelado. Si los reincidentes son castigados con severidad suficiente, la propensión a cometer una segunda infracción o infracciones ulteriores podría reducirse al mismo nivel que la propensión a cometer una primera infracción.

Si se admitieran libremente las pruebas anteriores a los delitos, y los jurados se inclinaran en gran medida a condenar a los infractores habituales independientemente de que las pruebas revelaran o no una culpa, se minaría la disuasión. El costo esperado del castigo de los infractores habituales bajaría porque ese costo es una función no sólo de la probabilidad del castigo en sí mismo, sino también, como vimos antes, de la diferencia existente entre la probabilidad del castigo en virtud de la culpa y la probabilidad del castigo dada la inocencia. (Sin embargo, un factor parcialmente compensatorio es el hecho de que habría un desincentivo adicional para *convertirse* en un infractor habitual.) El costo esperado del castigo de quienes delinquen por primera vez bajaría también. Los fiscales descubrirían que les resulta tanto más fácil obtener la condena de los infractores habituales, culpables o no, que se anularían sus incentivos para perseguir a quienes delinquen por primera vez, si suponemos que los fiscales operan con una restricción presupuestaria y quieren

maximizar las condenas ponderadas por la duración de la sentencia, sujetos a esa restricción.[18] Éstos son supuestos plausibles.

Hay una excepción general a la Regla 404 en los casos donde se acusa al indiciado de violación o de ataque sexual a niños y los delitos anteriores implican actos similares.[19] La excepción podría tener una justificación económica, con mayor claridad en el caso del abuso sexual. La mayoría de los individuos no se inclinan a abusar sexualmente de los niños. Por lo tanto, entre dos posibles corruptores de menores, sólo uno de los cuales tiene antecedentes de tales abusos, la historia establece una motivación que permite distinguir entre los dos sospechosos; y las pruebas anteriores a los delitos son admisibles para probar la motivación. Al revés de lo que ocurre con un corruptor de menores, un ladrón no tiene una *inclinación* por el robo, a menos que sea un cleptómano. El robo es sólo un instrumento de su deseo de tener dinero, y hay muchos instrumentos sustitutos. El hecho de que un indiciado acusado de robo cometiera un robo anterior no demuestra que "le guste" robar, de modo que no provee una motivación para su comisión del robo actual del que se le acusa.

La excepción más importante para excluir las pruebas del carácter, que se encuentra en la Regla 609, se refiere al uso de tales pruebas en los interrogatorios. Si un testigo ha sido condenado dentro de los últimos 10 años previos por un delito que implica el fraude u otro engaño, tal como la evasión de impuestos, la parte tiene derecho a utilizar esa condena en el interrogatorio a fin de "impugnar" (cuestionar la credibilidad de) su testimonio. Cualquier otra condena por un delito grave cometido dentro de los últimos 10 años puede utilizarse para este propósito, a menos que el juez obstaculice éste empleando la Regla 403, y aun una condena por delito grave cometido más de 20 años atrás puede ser utilizada si el juez (revirtiendo la presunción de la admisibilidad que aparece en la Regla 403) concluye que el valor probatorio de la condena para minar el testimonio del testigo supera su efecto perjudicial.

La justificación de la regla es que resulta improbable que una persona que ha infringido el derecho penal en el pasado tome en serio su juramento y testifique fielmente. Quizá sea así, ¿pero es *más* improbable que esta persona tome en serio su juramento que en el caso de quien delinque por primera vez y piensa que puede obtener una absolución si miente?[20] Tanto el reincidente como quien delinque por primera vez son delincuentes cuyo incentivo para mentir parece no relacionarse con el hecho de que hayan cometido un delito o más de uno. Es cierto que resulta más probable que un acusado mienta entre mayor sea el castigo que afronta si es condenado y que los reincidentes son castigados más severamente que quienes delinquen por primera vez. Pero el dato relevante es el castigo que afronta, no que sea un reincidente como tal.

[18] Véase Joel Schrag y Suzanne Scotchmer, "Crime and Prejudice: The Use of Character Evidence in Criminal Trials", 10 *J. Law. Econ. & Organization,* 319 (1994).

[19] Véase Fed. R. Evid., 413, 414; *cf.* Fed. R. Evid., 415.

[20] Si el acusado es inocente, presumiblemente prestará un testimonio veraz.

Al disminuir la probabilidad de que un infractor habitual que testifica sea absuelto, y al disuadir así a los infractores habituales de prestar testimonio —y con un efecto muy similar, ya que el jurado tiende a inferir la culpa del hecho de que el acusado no testifique (aunque se le instruya para que no tenga en cuenta esa abstención)—, la regla mina la disuasión de los delincuentes habituales, aunque una posible compensación es el aumento del costo de cometer la primera infracción.

XXIII.8. El privilegio

El derecho de una de las partes a excluir pruebas por razones normalmente no relacionadas con el interés en el apego a la verdad de los hechos se llama "privilegio". Examinemos el privilegio marital y el de abogado-cliente, posponiendo hasta el capítulo XXX la exclusión de las pruebas ilegalmente obtenidas y de las pruebas obtenidas en violación del privilegio contra la autoincriminación forzada.

Hay dos privilegios maritales. El privilegio *testimonial* se aplica a todas las comunicaciones entre esposos, ya se hayan realizado antes del matrimonio o en su transcurso, cuya revelación pública podría minar al matrimonio, pero sólo el cónyuge a quien se pide testificar puede invocar este privilegio. El privilegio de las *comunicaciones maritales* puede ser invocado por cualquiera de los cónyuges, pero se limita a las comunicaciones realizadas durante el matrimonio. Una confesión a un cónyuge en el sentido de que uno ha cometido un delito sería una prueba de culpabilidad muy fuerte, pero se excluye por temor de que se debilite el matrimonio al sembrar la desconfianza entre los cónyuges. Aun así, los privilegios maritales podrían inducir en efecto a casarse a algunas personas que no lo harían de otro modo, y es improbable que sean estables los matrimonios inducidos por un deseo de obtener ventajas probatorias. Y, al disminuir el costo del delito para los esposos, los privilegios estimulan a tales personas (aunque sin duda sólo muy débilmente) a cometer delitos; y el que un cónyuge cometa un delito es un acto muy desestabilizador para un matrimonio. Podría formularse un argumento más sólido en favor del privilegio si se limitara a los casos civiles.

Aun si los beneficios sociales de los privilegios maritales son escasos, los beneficios de su abolición serían también escasos, porque sería menos probable que los cónyuges se hicieran recíprocamente confesiones perjudiciales, de modo que la abolición no crearía una abundancia de pruebas valiosas. Esto contrasta con la regla de las medidas remediales discutida antes: si se permiten en un juicio las pruebas de reparaciones subsecuentes disminuiría el incentivo para hacer tales reparaciones, pero ello no lo eliminaría en medida importante, porque las reparaciones tienen beneficios en evitar la responsabilidad futura. Los beneficios de admitir una conducta ilícita contra nuestro cón-

yuge son menores que los beneficios de tomar medidas para evitar una responsabilidad futura que de otro modo sería muy probable; así, la abolición del privilegio marital podría hacer que tales confesiones se desvanecieran.

El privilegio de abogado-cliente, como sabemos, le prohíbe al primero divulgar, a menos que el segundo consienta en ello, lo que le haya dicho su cliente en el curso de su relación profesional. Concentrémonos en la aplicación del privilegio a las declaraciones hechas en el curso o la contemplación de la litigación, no en las declaraciones hechas en el curso de la búsqueda de asesoría legal en relación con los actos contemplados.[21] Como en el caso de una confesión hecha a nuestro cónyuge, una confesión hecha a nuestro abogado sería una prueba de culpabilidad muy fuerte. Pero se teme que el proceso acusatorio no funcionaría bien si las partes no pudieran hablar con sus abogados con entera confianza. Los clientes serían mucho más reservados acerca de lo que cuentan a sus abogados y, en consecuencia, estos últimos se verían en dificultades para darle seguimiento al caso del cliente. Otra consecuencia de la abolición del privilegio abogado-cliente sería la de inducir a los litigantes potenciales a invertir más en aprender por lo menos los rudimentos del derecho a fin de poder hablar con sus abogados con un riesgo mínimo de hacer confesiones comprometedoras. ¡La abrogación del privilegio podría aumentar así la inscripción en las escuelas de derecho! Además, al sentirse temerosos de obtener de sus clientes confesiones comprometedoras, los abogados podrían no obtener de ellos una información que el cliente no sabía que demostraría que su posición tenía méritos.[22]

XXIII.9. Los testigos expertos

La Regla 702 de las Reglas Federales de Evidencia permite que un testigo experto rinda testimonio acerca de cuestiones de su área de pericia "bajo la forma de una opinión". Una opinión es una inferencia obtenida a partir de una combinación de conocimiento de primera mano y conocimiento de antecedentes. Si vemos nubes negras y ofrecemos una opinión en el sentido de que está a punto de llover, la opinión representa la conjunción de la observación de las nubes con el conocimiento de antecedentes acerca de las señales del estado del tiempo. El extenso conocimiento de antecedentes que denota a una persona como experta en algún campo le permite ofrecer opiniones que serían irresponsablemente especulativas en boca de una persona que es lega.

Dada la complejidad técnica de muchas de las cuestiones que surgen en los juicios legales modernos, una gran dependencia respecto de los testigos

[21] Las diferentes consideraciones aplicables a los dos tipos de declaraciones se destacan en Louis Kaplow y Steven Shavell, "Legal Advice about Information to Present in Litigation: Its Effects and Social Desirability", 102 *Harv. L. Rev.,* 565 (1989).

[22] Este punto se destaca en Ronald J. Allen *et al.,* "A Positive Theory of the Attorney-Client Privilege and the Work Product Doctrine", 19 *J. Legal Stud.,* 359 (1990).

expertos parece ser la única alternativa a desplazarse hacia un sistema de tribunales especializados antes que (en gran medida) generalistas: un desplazamiento que tendría sus propios problemas.[23] No obstante, hay gran insatisfacción con el uso de testigos expertos. En virtud de que las partes respectivas les pagan, se teme que sean parciales en favor de quienquiera que los contrate ("sicarios" o "mentirosos pagados"), antes que testigos desinteresados y, por lo tanto, presumiblemente veraces. Por supuesto, empero, esto no los distingue claramente de varios otros tipos de testigos, en particular las partes mismas, a quienes se les prohibía antes toda testificación por esta razón. Y, en virtud de que la mayoría de los testigos expertos, al revés de lo que ocurre con la mayor parte de los testigos legos, son protagonistas reiterados, tienen un interés financiero en crear y preservar una reputación de ser honestos y competentes. Toda crítica judicial pública de un testigo experto (en una opinión publicada o no formalmente, o aun en la transcripción de un juicio o de otra audiencia) tiende a perjudicar la carrera del experto como un testigo, a veces fatalmente, porque es probable que la crítica surja en interrogatorios subsecuentes de este experto. Además, muchos testigos expertos son empleados por empresas de consultoría que tienen una reputación corporativa que puede verse perjudicada por los errores de cualquiera de sus empleados.

Esta observación puede utilizarse en defensa del testigo experto "profesional", a menudo ridiculizado. Entre más esperemos testificar como expertos en el futuro, mayor será nuestro interés en mantener una reputación inmaculada no provocando la crítica judicial, y, por lo tanto, será más creíble nuestro testimonio. De un testigo experto que ha testificado en docenas o centenares de casos sin suscitar la crítica judicial, y que espera testificar en muchos más casos, se tendrá muy probablemente la expectativa de que sea un testigo confiable. Puesto que, habiendo resistido los ataques de sus adversarios en sus casos anteriores, ha acumulado así un capital de reputación considerable que pondría en peligro al testificar deshonestamente en el caso actual. Aun así, es también el protagonista reiterado que tiene un incentivo para complacer a su cliente, de modo que sea contratado en el futuro. Dado que está prohibido dar a los testigos expertos un pago contingente, el testigo experto de una sola vez no tiene nada que perder *o ganar* si aporta un testimonio deshonesto o sesgado.

La Regla 706 de las Reglas Federales de Evidencia autoriza al juez a designar a sus propios testigos expertos, y ésta es una solución promisoria al sesgo que presentan los expertos designados por una de las partes y a la dificultad que tienen los jueces y, en especial, los jurados para resolver disputas entre expertos. Comoquiera que sea, son pocos los jueces que sacan provecho de esta autoridad. Un factor es el de la resistencia de los abogados. A éstos no les interesa que el juez o el jurado comprendan realmente el caso; lo que

[23] Véase también Richard A. Posner, *The Federal Courts: Challenge and Reform*, cap. 8 (1996).

les interesa es ganar, y, por consiguiente, contratan expertos (y les pagan muy bien) a quienes, según piensan, les resultarán convincentes los jueces o los jurados. En la renuencia del juez a designar expertos neutrales también desempeña un papel la aversión al esfuerzo. No sólo es que la labor de encontrar, entrevistar y conferir el puesto a expertos neutrales sea algo que toma tiempo, sino que, dado que los expertos nombrados por el tribunal no cuentan con un abogado que los guíe y los proteja, el juez debe desempeñar ese papel a menos que, como sucedió en un caso reciente, dé con un abogado dispuesto a representar a los expertos nombrados por el tribunal sin cobrar.[24] No obstante, es probable que esa modalidad obstaculice indebidamente el proceso. Y a los miembros del jurado les podrá causar perplejidad enterarse de que los propios testigos expertos del tribunal, aun cuando no sean partes involucradas en la demanda ni representantes de partes, de alguna manera necesitan a un abogado, quien presumiblemente los someterá de forma directa a un escrutinio al tiempo que, durante el interrogatorio, planteará objeciones a preguntas realizadas por los abogados de las partes que considere desorientadoras o de algún otro modo inapropiadas. Prescindir de un abogado para los neutrales significa más trabajo para el juez, a quien le corresponderá protegerlos contra los abogados de las partes que, deseosos de controlar el caso, pueden mediante una serie de preguntas agresivas socavar inapropiadamente la credibilidad de los expertos ante el jurado. Un factor final que da razón de la renuencia de los jueces a designar a sus propios expertos en los casos es la incongruencia que, según se percibe, guardan los expertos nombrados por el tribunal con las premisas del sistema acusatorio, el cual depende de las partes antes que del juez para obtener y presentar pruebas. Esta inquietud se puede mitigar si el juez da pasos encaminados a que los expertos de las partes expresen su consentimiento respecto al experto que él ha de nombrar.

Lecturas recomendadas

Dewatripont, Mathias y Jean Tirole, "Advocates", 107 *J. Pol. Econ.*, 1 (1999).

Harrison, Jeffrey L., "Reconceptualizing the Expert Witness: Social Costs, Current Controls and Proposed Responses", 18 *Yale J. on Reg.*, 253 (2001).

Hay, Bruce L., "Allocating the Burden of Proof", 72 *Ind. L. J.*, 651 (1997).

Kaplow, Louis, "On the Optimal Burden of Proof", 119 *J. Pol. Econ.*, 1104 (2011).

——— y Steven Shavell, "Legal Advice about Information to Present in Litigation: Its Effects and Social Desirability", 102 *Harv. L. Rev.*, 565 (1989).

[24] "Request and Notice regarding Dr. James Kearl, Filing 787", en Oracle America, Inc. *vs.* Google, Inc., 798 F. Supp. 2d, 1111 (N.D. Cal. 2011).

Katz, Avery, "Judicial Decisionmaking and Litigation Expenditure", 8 *Int'l Rev. Law & Econ.,* 127 (1988).
Landes, William M. y Richard A. Posner, "Harmless Error", 30 *J. Legal Stud.,* 161 (2001).
Milgrom, Paul y John Roberts, "Relying on the Information of Interested Parties", 17 *RAND J. Econ.,* 18 (1986).
Posner, Richard A., *Frontiers of Legal Theory*, caps. 11-12 (2001).
——— *Reflections on Judging,* cap. 9 (2013).
Sanchirico, Chris William, "Games, Information, and Evidence Production: With Application to English Legal History", 2 *Am. Law & Econ. Rev.,* 342 (2000).
Schrag, Joel y Suzanne Scotchmer, "Crime and Prejudice: The Use of Character Evidence in Criminal Trials", 10 *J. Law, Econ. & Organization,* 319 (1994).
Stein, Alex, "Inefficient Evidence" (Cardozo Legal Studies Faculty Research Paper, núm. 380, 11 de enero de 2013).
Tullock, Gordon, *Trials on Trial: The Pure Theory of Legal Procedure* (1980).

Problemas

1. Hay pruebas de que, en los casos de la responsabilidad de productos y la mala práctica médica, los juicios sin jurado favorecen mucho a los demandantes, mientras que los juicios por jurados favorecen mucho a los demandados.[25] La elección de un juicio de banquillo o sin jurado por parte del demandante, si tiene una argumentación fuerte, es explicable por la misma razón que la elección en el mismo sentido por parte del acusado de un delito si él tiene una posición fuerte. Lo inexplicable es que el demandado civil no escoja un juicio por jurados si tiene una posición débil y, por lo tanto, necesita un error del determinador de los hechos para ganar. Una respuesta posible se relaciona con el presunto mayor apego a la verdad de los hechos de los jueces. Decir que los jurados son menos apegados a la verdad precisa que los jueces equivale a decir que hay una varianza mayor de los resultados en los casos de jurados que en los de jueces. Un demandado que tenga una situación demasiado débil para tener alguna esperanza realista de escapar a la sentencia que lo considere responsable ante el demandante podría verse perjudicado por esta varianza. Supongamos que el intervalo de los posibles daños otorgados por un juez es de 10 000 a 100 000 dólares (con una media de 55 000), mientras que para un jurado en el mismo caso es de 0 a 110 000 dólares (la misma media). Entonces un demandado que no tiene ninguna probabilidad de persuadir al

[25] Kevin M. Clermont y Theodore Eisenberg, "Trial by Jury or Judge: Transcending Empiricism", 77 *Cornell L. Rev.,* 1124, 1162-1166 (1992).

jurado de que otorgue cero daños tiene más que perder si es enjuiciado por un jurado antes que por un juez.

¿Pero cómo podrá conciliarse esta conclusión con lo que dijimos antes, en XXII.3, en el sentido de que los demandados culpables en casos penales están en mejor situación si los juzga un jurado que si los juzga un juez?

2. En el siglo XIX se aplicaba la regla de que el demandante en un caso de negligencia tenía la carga de aportar la prueba de que no incurrió en una negligencia contribuyente, así como la prueba de la negligencia del demandado. En otras palabras, la cuestión de la negligencia contribuyente era parte de la acción del demandante antes que una defensa afirmativa. ¿Cuáles condiciones habrían hecho de esto la asignación económicamente justificada de la carga de la prueba?

3. Aun cuando la carga de la prueba civil estándar es la prueba aportada por una simple preponderancia de las evidencias, en algunos tipos de casos civiles —sobre todo los casos de fraude— se impone una prueba intermedia: la prueba por evidencias "claras y convincentes". ¿Puede encontrar una razón económica que respalde a esta imposición? ¿O una razón para que, en un caso de patentes, si se cuestiona la validez de la patente discutida en el juicio, la parte que hace el cuestionamiento deba probar el carácter inválido mediante evidencias claras y convincentes?

XXIV. LA APLICACIÓN DE LA LEY Y EL PROCESO ADMINISTRATIVO

XXIV.1. Aplicación pública de la ley contra aplicación privada: los dilemas

En los campos del derecho común tales como los cuasidelitos, los contratos y la propiedad, la aplicación de la ley —el proceso por el que se investigan las infracciones y se aplica al infractor una sanción legal— es, como su formulación, encomendada principalmente a personas privadas: los litigantes, sus abogados y diversos especialistas en la prueba o la investigación a quienes ellos contratan. Pero gran parte de la responsabilidad de la aplicación de la ley ha sido encomendada, concurrentemente con la aplicación privada o en forma exclusiva, a dependencias públicas. Este capítulo compara la aplicación pública con la privada, y también examina el tipo característico de responsable del cumplimiento público conocido como la agencia administrativa.

¿Para qué tener una aplicación pública? ¿No podría privatizarse por completo el cumplimiento de la ley?[1] Los individuos privados y las firmas legales investigarían las violaciones, aprehenderían a los violadores (incluidos los delincuentes) y echarían a andar procedimientos legales para corregir las violaciones, incluidos los procesos penales. Un responsable del cumplimiento privado estaría facultado, si tuviera éxito, para retener todos los resultados del juicio; por ejemplo, la multa pagada por el infractor condenado. Si el infractor estuviera a salvo del juicio, el Estado pagaría un premio al responsable del cumplimiento.

Esta propuesta resolvería el problema de la demanda colectiva como un instrumento de aplicación de origen privado, discutida en el capítulo XXII. Y, por radical que parezca, ello sería regresar a un procedimiento anterior de ejecución de la ley. La aplicación del derecho penal (como prácticamente de todo el resto) es casi enteramente privada en las sociedades primitivas y antiguas.[2] Durante muchos siglos en Inglaterra, el Parlamento y las autoridades municipales (así como las empresas privadas y los individuos) pagaban recompensas por la aprehensión y condena de los infractores. En el caso de las infracciones castigadas con multas, la multa se dividía entre la Corona y el responsable del cumplimiento. No había ministerios públicos, y la policía era

[1] Gary S. Becker y George J. Stigler, "Law Enforcement, Malfeasance and Compensation of Enforcers", 3 *J. Legal Stud.,* 1 (1974).

[2] Véanse, por ejemplo, Richard A. Posner, *The Economics of Justice*, caps. 5, 7-8 (1981); Michael Gagarin, *Early Greek Law* (1986); William Ian Miller, *Bloodtaking and Peacemaking: Feud, Law, and Society in Saga Iceland* (1990).

pública sólo de nombre.[3] Algunas huellas de este viejo sistema sobreviven en ciertos ámbitos del derecho estadunidense, entre los que destaca el juicio *qui tam,* que es una demanda levantada por un individuo de la esfera privada para reivindicar un derecho público —por ejemplo, un juicio entablado contra un contratista gubernamental por cobrarle excesivamente al gobierno—. Si la demanda es exitosa, el demandante logra quedarse con una participación de las ganancias, mientras que el resto se va a las arcas del gobierno.

Sin embargo, el cumplimiento privado impone ciertos costos que el público evita, lo que podría explicar la mezcla efectivamente existente de cumplimiento público y privado en el sistema legal actual. Supongamos que la multa por alguna infracción se encuentra sustancialmente por debajo del nivel óptimo y que la probabilidad de aprehensión y condena se encuentra sustancialmente por arriba de tal nivel, y que se aumenta la multa (f) tratando de aproximarse más al óptimo. Sólo si disminuye la probabilidad de aprehensión y condena (p), el aumento de la multa permitirá que se reduzcan los gastos en la aplicación de la ley sin ninguna disminución de la disuasión. Pero p aumentará con el cumplimiento privado. Aunque el aumento de la multa reducirá al principio el número de las infracciones al elevar el costo esperado de la infracción, también se incrementarán las ganancias de los responsables del cumplimiento que aprehenden a un infractor, y este último efecto podría conducir a un aumento del número absoluto de aprehensiones. Si así ocurre, está claro que el resultado será un aumento de p, porque p es simplemente la razón del número de aprehensiones al número de infractores. Pero, aun si el aumento de la multa tiene un efecto disuasivo tan grande que se traduce en menos aprehensiones, el rendimiento por aprehensión será mayor, y la competencia llevará a las empresas de la industria del cumplimiento a gastar más recursos que antes por cada aprehensión. Por lo tanto, aumentará la probabilidad de la aprehensión (p), lo que interferirá con los propósitos de la legislatura de aumentar la multa (ahorrar recursos reduciendo a p).

Una multa podría ser demasiado elevada antes que demasiado baja. Entonces, con el cumplimiento privado, la probabilidad de la aprehensión y condena tendría que ser demasiado baja, antes que demasiado elevada como en el ejemplo anterior. La razón es que, dado que $pf = C$ (los costos sociales del delito), si f es demasiado elevada, entonces, por asunto de simple aritmética, p debe ser demasiado baja para mantener la calidad de pf y C.

Con todo, si la legislatura redujera la multa a fin de corregir el problema, la probabilidad de aprehensión y condena disminuiría en lugar de aumentar (como debiera hacerlo para compensar la reducción de la multa), a medida

[3] Leon Radzinowicz, *A History of English Criminal Law and Its Administration from 1750* (1948). El Estado pagaba a la policía sólo sueldos nominales, y los policías buscaban en los regalos, las multas, etc., su compensación principal; en efecto, eran responsables del cumplimiento privado con licencia. Véase también Harold J. Krent, "Executive Control over Criminal Law Enforcement: Some Lessons from History", 38 *Am. U. L. Rev.,* 275, 290-295 (1989).

que los responsables privados del cumplimiento retiraran recursos de la industria en respuesta al precio menor de sus esfuerzos. Por lo tanto, podemos obtener un cumplimiento deficiente antes que excesivo; lo importante es que resulta difícil obtener, con el cumplimiento privado, la cantidad de cumplimiento correcto.

El problema desaparecería si la probabilidad óptima de la aprehensión y condena fuese de uno y los costos de la ejecución fuesen de cero, porque entonces la multa óptima sería igual a los costos sociales de la actividad ilegal; y, si tales costos aumentaran, la multa óptima aumentaría en la misma cantidad. Esto sería percibido (correctamente) por los responsables del cumplimiento como un desplazamiento hacia arriba de la curva de demanda que afrontan y tendría el efecto de aumentar los recursos destinados a la prevención del delito. No obstante, cuando la probabilidad de aprehensión y condena es menor a uno, la multa óptima es mayor a los costos sociales de la actividad ilegal, no como una señal de que debieran destinarse recursos adicionales a la prevención de la actividad, sino como un medio de minimización de tales recursos.

En cambio, en el caso del cumplimiento público, la multa no tiene que tomarse como una señal para invertir mayores recursos en la prevención del delito, ya que el responsable público del cumplimiento no está limitado a actuar como un maximizador privado de la ganancia. Igualmente, si la multa óptima es menor a la existente, una disminución de la multa actual no tiene que tomarse como una señal para invertir menos recursos en el cumplimiento.

En principio, la ejecución privada podría llevarse al nivel óptimo mediante impuestos (en el caso del cumplimiento excesivo con el que empezamos) o subsidios (en el caso del insuficiente). Consideremos el caso del cumplimiento excesivo. Por ejemplo, si la multa, *f,* fuera de 1 dólar, podría permitírsele al responsable privado del cumplimiento guardarse 90 centavos, mientras el resto iría a manos del gobierno bajo la forma de impuesto. Esto reduciría la demanda de cumplimiento privado hacia la izquierda sin disminuir *f*, reduciendo así el efecto disuasivo de la ley. Pero el impuesto introduciría una cuña entre lo que pagaran los infractores y lo que recibieran los responsables del cumplimiento, creando así oportunidades atractivas para el soborno; el infractor aprehendido y el responsables del cumplimiento estarían en mejor situación si negociaran un pago de transferencia privado que fuese menor a la multa legal, pero mayor a la multa menos el impuesto.

¿Podría el cumplimiento privado de la ley penal aumentar el número de los inocentes condenados por infracciones? Al responsable privado del cumplimiento se le paga supuestamente por infractor condenado, con independencia de la culpa o la inocencia efectivas del acusado. El responsable del cumplimiento puede aumentar de varias maneras su "pesca" y, por ende, su ingreso aumentando la oferta de "infractores". Puede fabricar una infracción; procesar a una persona inocente por una infracción que efectivamente ocurrió; o alentar a un individuo a cometer una infracción que no habría cometido sin tal estí-

mulo, y procesarlo luego por la infracción (véase antes VII.3).[4] Y, sabiendo que alguien está a punto de intentar cometer un delito, el responsable del cumplimiento puede esperar a que se cometa el delito y procesarlo luego, en lugar de aprehenderlo en la etapa de la preparación y procesarlo por una tentativa delictiva. El incentivo para la espera es obtener una compensación mayor, ya que la pena por el delito pleno será más severa que la pena por la tentativa.

Pero estos abusos ocurren también con el cumplimiento público: en efecto, las reglas contra la trampa, la supresión consciente de pruebas favorables para el acusado realizada por el responsable del cumplimiento, etc., se desarrollaron para impedir que los responsables públicos del cumplimiento utilicen estas prácticas. El responsable privado podría ser más sensible a los costos del procesamiento infructuoso. Esto podría llevarlo a detectar al inocente con mayor cuidado que el público, ya que es probable que los recursos destinados al procesamiento del inocente sean menos productivos que los recursos destinados al procesamiento del culpable. Y un responsable privado del cumplimiento podría obtener del acoso al inocente con acusaciones infundadas un rendimiento menor que el público. Al responsable privado del cumplimiento se le paga solamente si tiene éxito en el procesamiento (si bien existe la posibilidad de acosar a un demandado potencial a fin de recibir un pago con creces).

¿Es más probable que se cometa el error de procesar y condenar con el cumplimiento privado de la ley que con el público? No, es menos probable, en virtud de que las consecuencias financieras de no tener éxito en el procesamiento son más directas para el responsable del cumplimiento privado (a quien se remunera sobre la base de tarifa por pieza) que para el funcionario público asalariado.

XXIV.2. El cumplimiento público *versus* el cumplimiento privado: implicaciones positivas

El análisis de la sección anterior ayuda a explicar varias características del sistema legal:

1. Con pocas excepciones, existe un monopolio público —más precisamente, una serie de monopolios públicos— del cumplimiento del derecho penal. Es

[4] Una variante interesante de la trampa privada es la operación de escuelas para delincuentes por parte de responsables privados del cumplimiento en la Inglaterra del siglo XVIII. Tales escuelas podían ser racionales desde el punto de vista de maestros y estudiantes. El estudiante aprendía habilidades que disminuirían la probabilidad de su aprehensión por cualquier infracción particular que pudiera cometer. Esta disminución en el costo esperado del castigo del delito aumentaría la cantidad de la acción criminal, es decir, el acervo de infractores. Por lo tanto, desde el punto de vista de los responsables privados del cumplimiento, la operación de la escuela se asemejaría al llenado de un estanque por los pescadores.

cierto que el mismo acto es a menudo un delito y un cuasidelito, y entonces el cumplimiento privado es posible en principio. Pero si el infractor no responde a las pruebas, como ocurre con la mayoría de los infractores criminales, el remedio del cuasidelito resulta ineficaz y el responsable público del cumplimiento tiene un monopolio de hecho. En cambio, en áreas del derecho tales como los contratos y los cuasidelitos (excluidos los cuasidelitos que son también delitos), recaerá en el sector privado la carga principal del cumplimiento.

La diferencia relevante entre los delitos, por una parte, y los cuasidelitos y las violaciones de los contrato, por la otra, reside en el hecho de que, si se destinan muy pocos recursos a la aprehensión, esta probabilidad tiende a ser mucho menor a 1 en el primer caso y a aproximarse a 1 en el último caso. La víctima de una violación de contrato sabe quién es el promitente; la víctima de un accidente automovilístico conoce de ordinario la identidad del otro conductor,[5] pero la víctima de un robo con allanamiento conoce raramente la identidad del ladrón. Cuando no hay ninguna incertidumbre, no existe ningún riesgo de cumplimiento excesivo, siempre que los derechos de propiedad del cumplimiento no se asignen sobre la base de que se sirva primero al primero que llegue sino que se reserven para la víctima del cuasidelito o el incumplimiento de contrato. Si el daño marginal infligido por algún cuasidelito fuese de 9 dólares y los costos marginales de la aprehensión y condena fuesen de 1, de modo que $f = 10$, el responsable del cumplimiento que sea el primero en poner en juego su reclamación recibirá una renta de 9 dólares, ya que, si no fue una víctima, el componente de daño de la compensación sería pura ganancia en vez de una compensación. La oportunidad de obtener tales rentas originaría gastos en ejecución mayores a uno. Los derechos de exclusividad de la víctima eliminan esta fuente del dispendio. Incrementan el incentivo de la víctima para asistir en el procesamiento de la reclamación, aunque esto también ocasiona un riesgo de perjurio. Los derechos exclusivos de la víctima también disuaden a las víctimas potenciales de invertir excesivamente en la autoprotección. Recordemos que un objetivo de proporcionar compensación por los cuasidelitos y los incumplimientos de contrato consiste en inducir a la víctima potencial a hacer a un lado las precauciones no justificadas por el costo, con lo que la carga total de la protección se pone en las espaldas del victimario potencial (de quien se da por sentado que será capaz de evitar el daño a menor costo).

2. El costo del cumplimiento podría ser tan elevado en relación con el valor de la reclamación que no funcionaría el "mercado" de las reclamaciones legales si se observara una adhesión estricta al principio de que la víctima tiene el derecho exclusivo a la reclamación. Un buen ejemplo es el de una cons-

[5] Aunque esto se debe en parte al hecho de que las leyes penales y otros estatutos reguladores incluyan el licenciamiento de los conductores, el registro de los vehículos y el castigo por abandonar el escenario de un accidente: un ejemplo de la aplicación pública de la ley que trata de fortalecer el cumplimiento privado.

piración para fijar precios que impone un costo pequeño a cada uno de gran número de compradores. La demanda colectiva y la demanda por intermediario son instrumentos para solucionar este problema que ya hemos discutido. En efecto, los derechos de propiedad asignados normalmente a las víctimas de una violación reclamada son transferidos al abogado representante del grupo o del intermediario.

3. Los presupuestos de las dependencias públicas responsables del cumplimiento tienden a ser pequeños en relación con las ganancias potenciales del cumplimiento, tal como serían evaluadas por un responsable privado que maximiza sus ganancias. Por ejemplo, el Servicio de Impuestos Internos está operando a un nivel presupuestario en que el costo marginal del cumplimiento se encuentra muy por debajo del rendimiento marginal, medido (como lo haría un responsable privado) por la recaudación fiscal adicional que generarían los gastos adicionales en cumplimiento. El supuesto de una restricción presupuestaria sería poco realista si se aplicara a un responsable privado del cumplimiento porque este responsable podría financiar cualesquiera actividades de cumplimiento cuyo rendimiento esperado superara los costos esperados. Pero, si el Congreso aprobara fondos adicionales para una dependencia tal como el IRS que pudiera utilizar los fondos para aumentar el rendimiento neto del cumplimiento, podría generarse uno excesivo. El peligro de éste es particularmente agudo en el caso del IRS porque un incremento en la recaudación fiscal más allá de cierto punto produce costos sociales en vez de beneficios sociales incrementando el tamaño del gobierno más allá del nivel óptimo. (Asimismo, un cumplimiento más vigoroso de las leyes fiscales meramente desplaza la carga de la tributación de los contribuyentes no fraudulentos a los fraudulentos; la ganancia en eficiencia puede ser ligera, excepto en la medida en que los tramposos incurren en costos reales de evasión, generando así una pérdida de peso muerto.) La restricción presupuestaria de la dependencia pública opera así como operaría un impuesto sobre el cumplimiento privado destinado a disminuir el nivel de cumplimiento privado hasta el nivel socialmente óptimo.

4. Un monopolio público del cumplimiento permite, en efecto, que el responsable público anule leyes particulares, o cumplimientos particulares de la ley, simplemente absteniéndose de perseguir a los violadores. Tal anulación no sería un aspecto del cumplimiento privado; se aplicarían todas las leyes que generaran un rendimiento neto esperado positivo. ¿Es esto algo bueno o algo malo? Sabemos por el capítulo XXI que las normas jurídicas son a menudo demasiado incluyentes; los costos de la adaptación precisa de una norma a la conducta que se trata de prohibir son prohibitivos debido a las limitaciones inherentes de la presciencia y de las ambigüedades del lenguaje. No obstante, si se aplica al pie de la letra, una norma demasiado incluyente podría imponer grandes costos sociales; sería como el castigo de una persona inocente a fin de disminuir la probabilidad de absolver a un culpable. El "no cumplimiento"

discrecional es una técnica para disminuir los costos de la inclusión excesiva sin un aumento correspondiente de la inclusión deficiente (vacíos legales). La policía se hace de la vista gorda ante infracciones menores del código de tránsito vehicular; los inspectores de construcciones se desentienden de ciertas violaciones de las disposiciones del código de construcción cuyo castigo impediría la construcción de edificios nuevos en áreas urbanas; los controladores del tráfico aéreo permiten que las aerolíneas violen las regulaciones de seguridad demasiado estrictas que se aplican al espaciamiento del aterrizaje de los aviones o al despegue desde los aeropuertos.

Un monopolio público del cumplimiento es una condición necesaria, antes que suficiente, de la no ejecución discrecional. En principio, una dependencia pública podría ejecutar todas las leyes encomendadas a su administración. Pero en la práctica no puede hacerlo, dada la restricción presupuestaria antes mencionada.

5. El chantaje es la venta de derechos de información exclusivos a la persona que sería incriminada tras su revelación, y a primera vista parece un método eficiente para la ejecución privada de la ley (natural y positiva). El valor de la información para la persona chantajeada es igual al costo del castigo en que incurriría si la información se revelara y fuera castigada. Por lo tanto, esta persona estará dispuesta a pagar al chantajista hasta esa cantidad por los derechos exclusivos. La multa podría ser idéntica a lo que habría tenido que pagar si hubiese sido aprehendida y condenada por el delito que el chantajista ha descubierto. La criminalización del chantaje reduce el incentivo de las personas para buscar información que desacredite a otras porque no pueden vender los frutos de la búsqueda a la persona que más probablemente los compre a un precio alto.

Una respuesta que relaciona la prohibición del chantaje con el cumplimiento público de la ley es que, al limitar el mercado más lucrativo en el que pueden vender los coleccionistas privados la información desacreditadora (a saber, el de las personas a quienes la información desacredita), la prohibición disminuye el precio que las autoridades a cargo del cumplimiento legal tienen que pagar a estos coleccionistas; es un ejemplo de monopsonización. En congruencia con este punto, se permiten prácticas indistinguibles del chantaje en áreas del comportamiento donde la ley se aplica privada antes que públicamente. No se plantean objeciones (graves) contra una persona que recabe información acerca de las actividades adúlteras de su cónyuge, y que amenace con revelar esa información en un juicio de divorcio a fin de obtener una compensación máxima por el incumplimiento de las obligaciones maritales por parte del cónyuge ofensor. No se permite que un tercero chantajee al cónyuge ofensor; ello minaría la asignación del derecho exclusivo del cumplimiento de tales contratos a la víctima del incumplimiento.

Sin embargo, este análisis no explica la prohibición del chantaje cuando la información que el chantajista amenaza con revelar no se refiere a la acti-

vidad delictiva o ilegal por alguna otra razón, sino que es simplemente embarazosa. Aquí la explicación de la prohibición podría ser que el producto social del chantaje es escaso cuando el comportamiento o condición que vuelve vulnerable al chantaje a una persona es inelástico a las amenazas de revelación. Un ejemplo posible es el del comportamiento homosexual (pues los homosexuales han sido un blanco tradicional de los chantajistas), motivado en su mayor parte por un fuerte impulso instintivo, de modo que es improbable que se altere sustancialmente por la amenaza del chantaje. En tales casos, el efecto principal de la legalización del chantaje sería el incremento de los gastos de la detección y el ocultamiento de la homosexualidad antes que la modificación del comportamiento, y estos gastos se desperdiciarían desde un punto de vista social en general.

Se reportan sorprendentemente pocos casos de chantaje. La teoría de juegos sugiere que el chantaje puede ser raro, aun cuando el castigo no sea muy severo de ordinario y la detección pueda parecer difícil porque es improbable que la víctima del chantaje reporte el delito. El problema para el chantajista potencial es que le resultará difícil hacer creíble su amenaza de poner al descubierto el secreto de la víctima. Si el chantajista cumple la amenaza revelando el secreto de la víctima, será muy probable que ésta lo reporte a la policía porque ya no tiene nada que ganar cooperando con el chantajista. Sabiendo todo esto, el chantajista se dará cuenta de que su amenaza carece de credibilidad, sobre todo porque la penalización del chantaje les dificulta a los chantajistas establecer una reputación de hacer cumplir sus amenazas, una reputación en la que podrían haber invertido cumpliendo sus amenazas aun cuando eso fuese costoso para ellos.

Otro impedimento para el éxito del chantaje es que el chantajista, careciendo de un derecho de propiedad transferible sobre el secreto de la víctima, no puede hacer una promesa creíble de no reiterar sus demandas de dinero después de que la víctima cede ante su demanda inicial. Adviértase la analogía con la dificultad de monopolizar los bienes durables (véase antes IX.6).

6. El chantaje y el soborno son similares por cuanto el chantajista y el funcionario sobornado reciben un pago por dejar de aplicar la ley. Por lo tanto, sería de esperarse que en las áreas donde existe un monopolio público del cumplimiento se prohibiera el soborno tanto como el chantaje, mientras que se permitiera en las áreas donde no existe tal monopolio público. Y así lo observamos: el arreglo extrajudicial de un caso de cuasidelito o de contrato o antimonopolio privado es una forma de soborno enteramente legal, aunque en estos contextos no se usa ese término (¡excepto por los economistas!), dada su connotación peyorativa.

XXIV.3. Elección de casos por parte de las agencias públicas

La decisión de una agencia responsable del cumplimiento acerca de la concentración de sus recursos resulta interesante en vista de la posición monopólica que ocupan con frecuencia las agencias públicas en el cumplimiento de la ley. Exploraremos aquí esta cuestión bajo el supuesto de que la agencia actúa como un maximizador racional, comparando los rendimientos esperados con los costos esperados de usos alternativos de sus recursos. Este supuesto (revisado más adelante en XXIV.7) podría parecer incongruente con el hecho de que los organismos públicos de cumplimiento de la ley forman parte del proceso político. Pero no hay ninguna incongruencia. Las consideraciones políticas podrían afectar las ponderaciones empleadas por la agencia para determinar el rendimiento derivado de un triunfo en un tipo de caso particular: tales consideraciones podrían asignar un peso mayor al castigo del abandono de un cártel que al castigo de la participación en un cártel. Pero, una vez que se asignan estas ponderaciones y se determinan así las metas, la agencia podría tratar de asignar sus recursos con la mayor eficiencia posible para lograr sus metas.

Se critica a las agencia por dedicar recursos desproporcionados a casos triviales. El análisis económico sugiere que esta crítica es superficial. La utilidad esperada de un caso para la agencia es la ganancia que obtiene si triunfa, descontada por la probabilidad del triunfo. Supongamos que la agencia tiene que decidir cómo asignar un presupuesto fijo entre dos casos, A y B. A es el caso más importante. El triunfo en A aumentará en 100 unidades la utilidad de la agencia; el triunfo en B vale sólo 50 unidades; una derrota en ambos casos vale cero. Dado que la probabilidad de un resultado favorable depende en ambos casos de lo que gaste la agencia en el procesamiento, podría parecer que la dependencia debiera destinar a A el total o la mayor parte de sus recursos. Pero esto sería correcto sólo si los gastos de la agencia fuesen el único factor que afecta la probabilidad del resultado en cualquiera de los casos, lo que no ocurre. Los gastos del acusado son decisivos, al igual que la eficacia relativa de los gastos de las partes.

Si el caso es muy importante para el acusado, podría gastar mucho en su defensa. Entre más gaste, menos eficaces serán los gastos de la agencia en el litigio a menos que los aumente a fin de neutralizar los del acusado (véase antes XXII.11). En ambos casos la utilidad esperada del asunto para la agencia, una vez restados los costos de su procesamiento, será menor. Así pues, en igualdad de otras circunstancias, la agencia preferirá invertir recursos en un caso que sea poco relevante para el acusado. Esto podría ser importante porque, aunque las apuestas monetarias —que quizá sea lo único que le interesa al acusado— sean pequeñas, si la agencia gana el caso se sentará un precedente útil que aumentará la eficacia de sus gastos legales en casos futuros e impedirá por completo otras violaciones.

Hasta ahora hemos supuesto que está dado el número de casos que se presentan a la agencia, lo que desde luego no es cierto. A medida que una agencia conoce más casos de una clase particular, aumenta su utilidad total esperada, pero a una tasa decreciente porque los casos que son fáciles de ganar se volverán más difíciles de encontrar. Entre mayor sea la tasa de declinación de la probabilidad de éxito al aumentar el número de casos, será menor el planteamiento de casos. El universo de las violaciones menores es de ordinario mayor que el universo de las violaciones importantes; no "se le acaban" a uno tan pronto los casos fáciles de ganar. Ésta es otra razón para esperar que los casos pequeños predominen en la carga de trabajo de la agencia.

Sin embargo, existe todavía otra razón para esta preferencia, y constituye una señal de que hay un elemento de cumplimiento público que es ineficiente. Los directivos de una agencia por lo general planean quedarse en la oficina no más de un par de años, y lo mismo acontece con muchos miembros del personal; esta tendencia socava el incentivo para lanzarse en pos de proyectos que no aportan una compensación inmediata. Los abogados empleados por una agencia administrativa pueden preferir el manejo de pequeños casos en virtud de que eso les permitirá adquirir experiencia en materia de juicios de instrucción durante su breve tiempo de permanencia en la oficina, en tanto que un caso de gran envergadura podría no llegar al tribunal de instrucción sino hasta después de que se hayan ido: ellos sembrarían (es decir, llevarían a cabo investigación preparatoria y no poco tediosa) y les tocaría a otros cosechar (a sus sucesores, quienes en realidad instruirían el caso).

Se ha sostenido que los demandantes ganarán probablemente 50% de los casos litigados porque los errores de predicción del resultado —una condición necesaria para que ocurra la litigación, como hemos visto— son más probables en los casos difíciles.[6] Este argumento depende del supuesto más bien especial de que las partes tienen la misma información, muy apegada a los hechos, acerca del resultado de un caso si se va a juicio, lo que les permite arreglar todos los casos, excepto los más finamente equilibrados.[7] Dado que este supuesto no es válido a menudo, no es sorprendente que la hipótesis de 50% no esté bien corroborada.[8] Una razón independiente para que la mayoría de las agencias públicas tengan tasas de triunfo mucho mayores que 50% es que, al revés de los responsables privados del cumplimiento, operan bajo una restricción presupuestaria. Una agencia con una restricción presupuestaria estricta podría no plantear *ningún* caso difícil. (Esto se aplica a la mayoría de los casos penales hoy en día.) Por lo tanto, aunque la mayoría de sus casos se arregla-

[6] George L. Priest y Benjamin Klein, "The Selection of Disputes for Litigation", 13 *J. Legal Stud.,* 1 (1984); Joel Waldfogel, "Reconciling Asymmetric Information and Divergent Expectations Theories of Litigation", 41 *J. Law & Econ.,* 451 (1998).

[7] Véase Steven Shavell, "Any Frequency of Plaintiff Victory at Trials Is Possible", 25 *J. Legal Stud.,* 493 (1996).

[8] Véase *ibid.* en 500-501.

rán extrajudicialmente, sus juicios de instrucción se seleccionarán todavía entre una población de casos que se inclinan claramente hacia un lado.

Un análisis completo de la regulación pública integraría el análisis del cumplimiento público (este capítulo) con el análisis del castigo (capítulo VII). Para ilustrar, consideremos el efecto de un aumento de la progresividad sobre el cumplimiento de las leyes del impuesto al ingreso.[9] Podría parecer obvio que el cumplimiento de los contribuyentes de altos ingresos disminuiría y que el de los contribuyentes de bajos ingresos aumentaría, de modo que un IRS racional reasignaría recursos de auditoría de los contribuyentes de ingresos bajos a los de ingresos altos. Pero no ocurre así necesariamente. Es cierto que el aumento de la tasa impositiva marginal del contribuyente de altos ingresos incrementa la ganancia de la declaración con ingresos subestimados, pero también incrementa la ganancia que obtiene el IRS al auditar a los contribuyentes de altos ingresos. El último efecto podría predominar por dos razones. Primero, si los contribuyentes sienten en general aversión por el riesgo, la ganancia incrementada por la subestimación de sus ingresos en la declaración y la correspondiente pérdida incrementada si se descubre la falta vuelven a la declaración de ingresos subestimados una estrategia más riesgosa y, por lo tanto (en igualdad de las demás circunstancias), menos atractiva. Segundo, e independientemente de la actitud asumida hacia el riesgo, la mayor pérdida esperada de ser descubierto disminuye la mayor ganancia esperada de la declaración incompleta. El contribuyente racional sabe que el IRS tiene una ganancia esperada de su auditoría mayor que la suya por efecto de la declaración incompleta de sus ingresos, lo que lo llevará a ocultar menos ingresos, lo que a su vez permitirá que el IRS reasigne recursos de auditoría a los contribuyentes de bajos ingresos, de quienes puede demostrarse mediante una inversión sencilla del análisis anterior que tienen un incentivo mayor que antes para presentar declaraciones incompletas.

Una interrogante importante acerca del cumplimiento tributario consiste en saber si el IRS debiera revelar sus criterios para la auditoría de las declaraciones de impuestos. La respuesta intuitiva es que no debiera hacerlo porque hacerlo alentaría a los contribuyentes a declarar deficientemente hasta el punto en el que se desataría una auditoría. Pero esto podría ser incorrecto. Un contribuyente podría inferir del secretismo del IRS que los criterios son más laxos de lo que son en realidad, de modo que podría declarar menos ingresos. Sería particularmente probable que el contribuyente obtuviera tal inferencia si las dependencias que alcanzan una tasa de auditoría muy elevada proclaman el hecho. Esto es análogo al escenario del registro de calificaciones del estudiante, en el cual el incentivo de algunos miembros del grupo para divulgar

[9] Véase Michael J. Graetz, Jennifer F. Reinganum y Louis L. Wilde, "The Tax Compliance Game: Toward an Interactive Theory of Law Enforcement", 2 *J. Law, Econ. & Organization,* 1 (1986).

información privada finalmente obliga a todos los demás miembros del grupo a revelar el dato.

Un método controvertido para aplicar la tributación es el de la "amnistía fiscal". En efecto, al evasor fiscal se le vende una inmunidad al castigo a cambio del pago que debe de impuestos atrasados. La amnistía es atractiva para el gobierno porque eleva la recaudación, y lo es para el evasor fiscal porque le permite comprar su derecho a evadir el riesgo de ser procesado en caso de que lo pesquen con las manos en la masa. Se trata de un intercambio mutuamente benéfico. La objeción a las amnistías es que incrementan el incentivo para cometer ese delito en el futuro al ofrecer la perspectiva de futuras amnistías. La objeción es superficial. El gobierno (si es sensato) habrá de sopesar el trueque habido entre la ganancia obtenida en recaudaciones merced a la amnistía y la futura pérdida de ingresos fiscales que probablemente cause el proyecto de futuras amnistías, y establecerá el "precio" de la amnistía a un nivel que maximice la ganancia neta en recaudaciones. Por ejemplo, si calcula que la perspectiva de amnistías futuras habrá de conducir a evasiones fiscales significativamente más cuantiosas en el porvenir, puede condicionar la amnistía a que el evasor fiscal pague no sencillamente los impuestos atrasados que debe sino también una multa sustancial.

El IRS ha empleado amnistías fiscales con bastante frecuencia en años recientes;[10] esto aporta cierta evidencia de que se trata de instrumentos eficientes para aumentar la recaudación fiscal.[11]

El supuesto hasta ahora sostenido en este capítulo de que las agencias se comportan como maximizadores racionales (dadas cualesquiera ponderaciones distributivas asignadas por la legislatura a los resultados del cumplimiento de la dependencia) sugiere una razón económica para proporcionar asesoría pública a los acusados penales indigentes. El éxito de una agencia fiscalizadora podría depender de que gane el mayor número posible de casos (ponderados por su importancia) con el menor gasto posible. Dado que pue-

[10] Véase Lynnley Browning, "14,700 Disclosed Offshore Accounts", *N.Y. Times,* 17 de noviembre de 2009, www.nytimes.com/2009/l1/18/business/global/18irs.html?_r.=0; Internal Revenue Service, "IRS Says Offshore Effort Tops $5 Billion, Announces New Details on the Voluntary Disclosure Program and Closing of Offshore Loophole", 26 de junio de 2012, www.irs.gov/uac/IRS-Says-Offshore-Effort-Tops-$5-Billion,-Announces-New-Details-on-the-Voluntary-Disclosure-Program-and-Closing-of-Offshore-Loophole.

[11] Para más análisis económicos de las amnistías fiscales, véase James Andreoni, "The Desirability of a Permanent Tax Amnesty", 45 *J. Pub. Econ.,* 143 (1991); Arun S. Malik y Robert M. Schwab. "The Economics of Tax Amnesties", 46 *J. Pub. Econ.,* 24 (1991); Peter Stella, "An Economic Analysis of Tax Amnesties", 46 *J. Pub. Econ.,* 383, 397-398 (1991). Al momento de escribir estas líneas se está dando una discusión extensa acerca de una posible amnistía a los inmigrantes ilegales, quienes conforman una suma de 11 millones en los Estados Unidos. La parte capciosa del asunto consiste en lograr que los términos de la amnistía sean lo suficientemente dolorosos (por ejemplo, al imponer un periodo de espera sustancial para obtener la ciudadanía estadunidense) para que de todos modos sea disuadida la inmigración ilegal en el futuro, como no lo sería si los inmigrantes ilegales del día de mañana estuvieran confiados en recibir una amnistía.

de resultar más barato condenar a un acusado que no esté representado por un abogado que a uno que sí lo esté, aunque el primero sea inocente y el segundo culpable la agencia podría tener un incentivo para procesar a gente inocente que sea indigente. El resultado sería la imposición de costos de castigo socialmente improductivos, la disminución del efecto disuasivo del castigo penal y la desviación de recursos de persecución de las áreas donde podrían ser empleados más productivamente desde el punto de vista de la sociedad. Por lo tanto, el derecho a la asesoría puede defenderse como un método de impartición de los incentivos correctos a los responsables del cumplimiento de la ley.[12]

En virtud de que la producción de una agencia pública es a menudo imposible de medir, los funcionarios y los legisladores que las monitorean podrían preferir una medida de la producción que, por burdamente que se correlacione con los beneficios sociales, pueda determinarse por lo menos con una certeza considerable. Así, dado que es mucho más fácil determinar el número de inmigrantes ilegales aprehendidos cada año que la del número de personas que logran inmigrar ilegalmente cada año, la Patrulla Fronteriza de los Estados Unidos basa sus peticiones de presupuesto en el primer número antes que en el último.[13] Esto genera un incentivo perverso para permitir que los ilegales entren en bandadas, ya que entre más logren cruzar la frontera más podrán ser aprehendidos a bajo costo. Otro ejemplo lo brinda la poco cómoda combinación de investigación penal e inteligencia de seguridad nacional en el Buró Federal de Investigaciones (FBI).[14] La producción de investigaciones penales se puede medir objetivamente en términos de los arrestos ponderados por las condenas ponderadas por la severidad de las sentencias y la cantidad de propiedad recuperada. La producción de inteligencia es difícil de medir objetivamente debido a que no hay ningún buen punto de referencia para determinar si el servicio de inteligencia está llevando a cabo un buen trabajo. Naturalmente, una agencia favorecerá aquella actividad en la que pueda presentar al público y a los comités de asignaciones de las legislaturas pruebas sólidas de que ha cosechado el éxito, y, de este modo, uno espera que el Buró dé mayor relieve a la ejecución de la ley penal que a la inteligencia.

A lo largo de esta sección, hemos dado tácitamente por sentado que la función de utilidad relevante es la de la agencia. El supuesto relega los costos de agencia (en el sentido de principal-agente), que probablemente son más grandes en el caso de una agencia pública que en el de una empresa de negocios. La producción de una compañía se puede cuantificar con facilidad, lo que

[12] Más adelante, en XXV.3, se considera otra posible justificación económica del derecho.

[13] Véase Alberto Dávila, José A. Pagán y Montserrat Viladrich Grau, "Immigration Reform, the INS, and the Distribution of Interior and Border Enforcement Resources", 99 *Pub. Choice,* 327 (1999).

[14] Véase Luis Garicano y Richard A. Posner, "Intelligence Failures: An Organizational Economics Perspective", *J. Econ. Perspectives*, otoño de 2005, 151.

ofrece un punto de partida para evaluar el desempeño de un empleado, aun cuando frecuentemente la aportación precisa de un empleado en particular no se puede determinar con ninguna precisión. El problema de la evaluación es más arduo en el caso de los empleados públicos porque rara vez se puede cuantificar la producción de las agencias públicas. Cuando la producción de un empleado resulta inconmensurable, prevalece la tendencia a fundar la contratación y la promoción en las entradas o materias primas que aporte, abrigando la esperanza de que la calidad del insumo y la calidad del producto acabado se correlacionen positivamente. De ahí la importancia que se asigna a las referencias que comprueban la valía de un candidato judicial, las cuales pueden ser comparadas y ordenadas según una jerarquía. El caso del FBI ilustra con creces la forma en que un intento de valorar la producción o producto terminado se puede sustituir por una evaluación objetiva de entradas. El producto final que desea obtener el FBI en las investigaciones penales no se traduce en el número de arrestos, condenas y sentencias, sino en la reducción de la incidencia de delitos a escala federal. Los arrestos y demás son sencillamente insumos dirigidos al logro de una reducción en los delitos, pero resultan mucho más fáciles de medir y existe una correlación negativa entre la ejecución de la ley y la actividad delictiva.

Otra fuente común de ineficiencia en las agencias públicas es la de las "guerras territoriales". Se encuentra igualmente en las empresas privadas, pero predomina de manera particular en el gobierno. El porqué de estos acontecimientos es la raquítica definición de derechos de propiedad y la dificultad consecuente que hay para zanjar disputas de modo transaccional. Supongamos que una agencia cuenta en su edificio con un espacio extra que otra agencia necesita urgentemente. La primera agencia se mostrará en extremo renuente a conceder el espacio, en virtud de que es improbable que obtenga una retribución a cambio de ello. Todo lo que conseguirá por su generosidad será alcanzar una reputación de debilidad que podría alentar a otras agencias a invadir su territorio.[15]

XXIV.4. La regulación de los tribunales *versus* la regulación de las agencias administrativas

Las agencias administrativas están especializadas, lo que propicia el desarrollo de habilidades consumadas en temas técnicos (siendo buenos ejemplos el de los departamentos de seguridad en el tráfico que prescriben límites de velocidad y el de la Administración de Alimentos y Medicamentos [FDA] que regula los productos farmacéuticos). Por lo general reclutan a un personal numeroso

[15] *Cf.* Steven D. Levitt y Sudhir Alladi Venkatesh, "An Economic Analysis of a Drug-Selling Gang's Finances", 115 *Q. J. Econ*., 755, 780-781 (2000).

y poseen facultades de característica flexibilidad: a menudo tienen autorización para utilizar tanto la regulación *ex ante* como la *ex post* (elaboración de reglas y adjudicación). Se encuentran menos maniatadas por el precedente que los tribunales. Los miembros de las agencias gozan de más legitimidad política que los jueces porque no disfrutan de una titularidad de por vida y, por ello, tienen menos necesidad de evitar que los clasifiquen como "activistas" y de demostrar que observan la continuidad con arreglos políticos establecidos en el pasado. Los jueces son renuentes a innovar, o al menos a dar muestras de su capacidad innovadora, so riesgo de ser acusados de atravesar la frontera que separa el cumplimiento de la ley de la elaboración de la ley, siendo esta última área considerada en la jurisprudencia ortodoxa y la teoría política como una función legislativa antes que judicial. Los jueces siempre están negando que formulen las políticas… algo que las agencias llevan a cabo sin dar disculpas. En verdad, la diferencia es meramente de grado.

Al mismo tiempo, las agencias administrativas se hallan sujetas a presiones de los grupos de interés bastante más intensas que los tribunales. Los jefes de las agencias son gente designada por motivos políticos y su trabajo es objeto del monitoreo que realizan los comités del Congreso. El hecho de que los miembros de las agencias estén especializados, y de que se encuentren menos aislados del proceso político que los jueces, los hace blanco de los grupos de intereses particulares en busca de esferas de influencia; de ahí el uso del término "captura del regulador" (véase XXIV.6). Con frecuencia, las decisiones de las agencias reguladoras han sido anticompetitivas. Los grupos de interés tratan de influenciar a las agencias con la finalidad de mantenerse a buen resguardo de la competencia, por ejemplo, poniendo obstáculos al ingreso de nuevos grupos. La ejecución de políticas regulatorias válidas muchas veces se ve frustrada por la agencia de los reguladores respecto de la información que suministran las entidades reguladoras y por los incentivos perversos que crea el comportamiento de "puerta giratoria". El personal nutrido de la mayor parte de las agencias reguladoras da lugar a los típicos problemas de costos de agencia padecidos por las burocracias que no se supeditan a la competencia del mercado. Y la regulación es en realidad una regulación dual de agencia-tribunal, porque los ordenamientos de las agencias son apelables ante un tribunal. La regulación puede, como en el caso de los beneficios de seguridad social por discapacidad, engendrar cuatro niveles de adjudicación: un juez de derecho administrativo en la Administración de Seguridad Social, una revisión a cargo de un cuerpo de apelaciones en el seno de la administración, la revisión a cargo de un tribunal distrital a escala federal y la apelación de ese tribunal ante un tribunal federal de apelaciones. No es posible hacer economías en una regulación provista de semejante multiplicidad de niveles.

Los tribunales son relativamente inmunes a las presiones de los grupos de interés (al menos los tribunales federales, cuyos jueces tienen una titularidad ganada, y algunas cortes estatales); presentan un carácter no burocrático; están

descentralizados, y son semiprivatizados (a causa del papel protagónico que desempeñan los abogados de los litigantes). Ponen al alcance la perspectiva propia de un forastero (un recién llegado de otros campos del saber) sobre asuntos que los reguladores, víctimas del padecimiento denominado "visión de túnel", podrían conducir a la ruina. Los jueces están también menos imbuidos del sentido de haber sido llamados a una misión que los reguladores. Siendo generalistas, y al provenir de ámbitos profesionales y contar con antecedentes personales diversos, es menos probable que se identifiquen con políticas determinadas y, por consiguiente, aportan un enfoque más equilibrado de los problemas que el que manejan los reguladores, comprometidos con una directriz política en particular. (Esto conlleva una argumentación contra las judicaturas especializadas, que son comunes en los sistemas legales de Europa continental.) Si se estableciera una agencia para erradicar el narcotráfico y se le concediera la autoridad para procesar a los violadores de las leyes contra los estupefacientes, les prestaría escasa atención a las salvaguardas procedimentales para los infractores acusados. Los jueces federales gozan de mayor prestigio, de mejores condiciones laborales, de un salario y (en particular) unos beneficios más elevados, de más seguridad laboral y de bastante más autonomía (particularmente evidenciada en su protección contra las presiones políticas y de los grupos de interés) que los reguladores, y todas estas ventajas traen por resultado una calidad promedio más alta de los individuos designados para puestos en el poder judicial que en el sector regulatorio.

Aun así, la falta de conocimientos especializados de los jueces, la proporción reducida de su personal, el coto impuesto a sus recursos para la investigación, los procedimientos llenos de vericuetos y hasta cierto punto anticuados que deben seguir, el compromiso que deben asumir de cara al incremento en la formulación de reglas y, en fin, la demora en que incurren al responder a problemas sociales de índole severa —los tribunales no pueden actuar hasta que llega a la sala un caso, lo que a menudo ocurre mucho después de haberse iniciado la actividad que dio lugar al caso— constituyen impedimentos para una regulación efectiva, especialmente de temas técnicos. Estos problemas se agravan debido al uso excesivo —peculiar según los estándares mundiales— de jurados en casos civiles. Cuando se encarga a las cortes ventilar asuntos técnicos, tales como los problemas concernientes a la mala práctica médica, los defectos en el diseño de los productos y las patentes farmacéuticas o de software, los resultados son frecuentemente todo menos satisfactorios. No obstante, el rendimiento deficiente del Tribunal de Apelaciones para el Circuito Federal, el cual tiene un carácter semiespecializado y posee el monopolio de apelaciones sobre patentes, sugiere que la especialización de los tribunales o las agencias no es ninguna panacea incluso en campos de alto nivel técnico.[16]

[16] Véase William M. Landes y Richard A. Posner, "An Empirical Analysis of the Patent Court", 71 *U. Chi. L. Rev.,* 111 (2004).

El auge de la agencia administrativa federal, que comenzó cuando fue creada la Comisión Interestatal de Comercio (ICC) en 1887, reflejaba el anhelo de aumentar el conocimiento consumado en la regulación y de servir de contrapeso a lo que, según se creía ampliamente y hasta cierto punto de manera correcta, era la obstinada resistencia judicial a las políticas modernas de bienestar social. Pero también reflejaba el anhelo de proporcionar, ante el espectro de las que se percibían como acciones delictivas, remedios más módicos y expeditos que administrarían los servidores civiles. Las ganancias cosechadas fueron un grado de pericia, de resolución expedita y de flexibilidad en el procedimiento y el remedio; las pérdidas sufridas fueron la habilidad superior de la mayoría de los jueces (al menos los jueces federales) en comparación con la de los administradores de las agencias, su sensibilidad más aguzada en lo tocante a los valores en que se cifran las reglas de la ley y la energía e iniciativa de las personas y compañías en el ámbito privado a las que afectó la regulación.

Un sistema puro de regulación sería sólo una regulación administrativa o exclusivamente una litigación; un sistema mixto combina las dos modalidades de control. No existe prácticamente ningún sistema regulatorio puro, ya que la mayor parte de las decisiones regulatorias tomadas por las agencias administrativas están sujetas a la revisión judicial. Ésa es una forma de intervención judicial más limitada que un procedimiento que empieza en el tribunal en vez de en una agencia. Con todo, por otra parte, resulta común suministrar paralelamente remedios administrativos y judiciales para los mismos daños, como en el caso del antimonopólico, donde uno puede plantear una demanda en la corte federal o levantar una queja ante la Comisión Federal de Comercio, a quien corresponde decidir si se ha de instituir un procedimiento administrativo, el cual será finalmente objeto de una revisión judicial. Se encuentra una clase diferente de regulación dual en lo relativo a la discriminación en el empleo. Una persona que se queja de discriminación en un acto violatorio de la ley federal puede levantar una demanda en la corte federal. De modo alternativo, la Comisión para la Igualdad de Oportunidades en el Empleo (EEOC) puede plantear una demanda en la corte federal en representación de la persona; en ningún caso recae esa adjudicación en la comisión misma, a diferencia de lo sucedido en el ejemplo del procedimiento antimonopólico.

Aunque los tribunales federales se niegan a emitir, como sabemos, opiniones consultivas, las agencias federales, como el Servicio de Impuestos Internos y la Comisión de Bolsa y Valores, ofrecen una buena porción de asesoría, en caso de petición, a las personas sometidas a su autoridad regulatoria. ¿Cuál es la razón de la diferencia en la disposición a prestar consejo?[17] Los tribunales se fundan en una disputa concreta para poner a las personas al tanto de los hechos esenciales; y la mayoría de los tribunales estadunidenses tienen una

[17] Véase William M. Landes y Richard A. Posner, "The Economics of Anticipatory Adjudication", 23 *J. Legal Stud.,* 683 (1994).

jurisdicción general en lugar de una especializada. Carecen de una buena base probatoria para fungir de asesores y, asimismo, se verían inundados por peticiones de asesoría, y bien sabemos que la decisión de expandir un sistema de tribunales se topa rápidamente con severas deseconomías de escala. Las dependencias tienen otras fuentes de información además de las disputas concretas, y disponen de una jurisdicción circunscrita que limita el número potencial de peticiones de asesoría. Cuando éstas las hace una agencia administrativa, se trata, además, de solicitudes autolimitantes, ya que una petición puede poner al descubierto información acerca del solicitante que suscite el escrutinio regulatorio. Más importante aún, si la agencia le advierte al solicitante que un determinado curso de acción es ilegal y el solicitante se inclina por él de todos modos, la agencia, a fin de mantener la credibilidad de su función de asesora, *tendrá* que procesar el caso, en tanto que si no se hubiera buscado ningún consejo la agencia podría resolver que era mejor ignorar el asunto. Este efecto de disuadir a alguien de buscar a quien lo asesore se halla ausente si una corte es el cuerpo oficial cuyo consejo se está intentando conseguir, debido a que una corte no emprende la prosecución.

Lo más cercano a un sistema puro de regulación administrativa es un sistema en el que el acatamiento de una regla u orden regulatorios impide que se dé en forma subsecuente un juicio (acción preventiva). La industria financiera presenta lo que se aproxima más a un régimen regulatorio puro. La mala práctica médica se aproxima a un sistema puro de litigación, salvo que existe cierta regulación de las prácticas hospitalarias y médicas, y, por supuesto, tiene que darse la acreditación de médicos y de otros expertos en el cuidado de la salud. La regulación antimonopólica se acerca también a un sistema puro de litigación, pese a los lineamientos de fusión publicados por el Departamento de Justicia y la FTC. Los lineamientos proveen de una base para la toma de resoluciones firmes por parte de las agencias sobre la conveniencia de aprobar las fusiones propuestas, pero no son vinculantes. Paralelamente, se encuentran los lineamientos federales para las sentencias: se trata de directrices administrativas que influyen en el dictado de sentencias penales a nivel federal pero que no son vinculantes.

Un sistema puro es más barato; más simple; opera con mucha mayor rapidez, y proporciona una guía más sólida. La necesidad de velocidad, bien ilustrada por la respuesta de la Reserva Federal y del Departamento del Tesoro al súbito colapso financiero en septiembre de 2008, puede ser una razón contundente para implementar un sistema regulatorio en el que los tribunales desempeñen un papel pequeño o nulo. Así pues, mientras las firmas que se encuentran en declive normalmente están sujetas a la liquidación o a la reorganización en una corte de quiebras (parte del sistema de tribunales federales), los bancos comerciales que se vienen abajo hallan su "resolución" en los procedimientos administrativos emitidos por parte de la Corporación Federal de Seguro de Depósitos (FDIC).

Cuando el factor del tiempo no es esencial, un sistema puro de regulación administrativa tiene muchas desventajas. Para empezar, incrementa a ojos de los grupos de interés los incentivos para controlar a las agencias regulatorias y, por lo tanto, su probabilidad de hacerlo; un grupo de interés sólo tiene que "comprar" a una agencia, y no también a los tribunales… y los tribunales federales son muy difíciles de "comprar". Asimismo, un sistema puro de cualquier especie (agencia o tribunal) sacrifica las complementariedades, porque los tribunales y las agencias son complementos a la vez que sustitutos. La acción de la agencia sujeta a la revisión judicial mezcla las perspectivas especialistas con las generalistas y la política encaminada a la misión con la sensibilidad ante los factores concernientes a la regla del derecho.

Tal como se acaba de mencionar, un sistema mixto explota las complementariedades entre las agencias y los tribunales. Los jueces prontos a emitir sentencia se acoplan bien a los lineamientos sentenciadores; los fallos antimonopólicos se acoplan a los lineamientos del Departamento de Justicia o de la FTC en materia de fusiones, y los jueces revisan los ordenamientos de las agencias administrativas para ver si se acatan los estatutos y los principios de procedimientos justos, temas con los que ellos están más familiarizados y a los que suelen dar un peso apropiado con mayor escrúpulo que el que están capacitados para darles los administradores, más orientados hacia el sentido de misión. Un sistema mixto es menos vulnerable a la dominación ejercida por los grupos de interés en virtud de que en él dichos grupos tienen que "comprar" tanto a la agencia como a los tribunales. (Éste es un argumento tradicional a favor de la instrucción realizada por el jurado en vez de por el juez: es más difícil sobornar a 12 "jueces" que a uno.) El sistema mixto también suministra un respaldo o una capacidad regulatoria a prueba de fallas. En el caso de la seguridad farmacéutica, por ejemplo, cuando la Administración de Alimentos y Medicamentos no logra prevenir la venta de un producto farmacéutico que es riesgoso, la ley de cuasidelitos correspondiente a la responsabilidad de productos ofrece una alternativa: la del control *ex post* sobre su venta.

XXIV.5. La superposición de las agencias

A menudo más de una agencia regula la misma actividad. Tanto el Departamento de Justicia como la Comisión Federal de Comercio aplican las leyes federales antimonopólicas, y, además, los fiscales generales estatales llevan a cumplimiento las leyes antimonopólicas estatales, que se ajustan al molde de las leyes federales y son aplicables a muchas de las mismas empresas. También es posible inducir a los juicios privados a que cumplan las leyes antimonopólicas federales y estatales. Para hacer más complicado el cuadro, los fiscales generales estatales pueden ventilar los juicios antimonopólicos federales en representación de sus estados. Y, en el área de regulación de la segu-

ridad, por ejemplo, es común que los estados impongan estándares de seguridad más rigurosos que los reguladores federales, aun cuando se otorgará a la regulación federal la facultad de tomar la delantera sobre la regulación estatal que la contradiga, de manera que, en el caso de ser ésta aplicable, impondría obligaciones incongruentes en las empresas reguladas.

La competencia regulatoria aumenta las probabilidades de que una violación sea detectada y castigada (en parte volviendo más difícil que las empresas reguladas se coludan con sus reguladores[18]), propicia la experimentación con enfoques diversos y actúa como una especie de seguro parcial contra el fracaso regulatorio.[19] Pero también tiende a crear redundancias dispendiosas[20] y así incrementa los costos de cumplimiento para las firmas sometidas a los regímenes regulatorios duales o múltiples.

Otra objeción al traslape regulatorio es que da nacimiento al "arbitraje regulatorio": la práctica sostenida por las compañías de configurar sus negocios de tal manera que los colocan bajo la jurisdicción regulatoria de una agencia que tiene probabilidades de favorecer a la firma, quizá debido a que la agencia se sostiene con cuotas de las firmas que regula, y, por lo tanto, para incrementar su presupuesto, busca tentar a las firmas para que se coloquen bajo su jurisdicción mediante la promesa implícita de una regulación ligera. De esta manera, un banco podría decidir lanzarse en búsqueda de una Constitución estatal en vez de federal por pensar que el comisionado bancario estatal se mostraría más tolerante ante las políticas de préstamo del banco que un regulador bancario federal, y, además, el comisionado podría dar la bienvenida al recién llegado a causa del efecto que tendría sobre su presupuesto contar con un nuevo "cliente" atento a pagar sus cuotas. Adviértase el estrecho paralelo con la compra de foros legales (o búsqueda de los tribunales más propensos a emitir fallos favorables) en la litigación. El arbitraje regulatorio fue un factor en el colapso financiero de septiembre de 2008; muchos bancos y otras empresas financieras se habían ubicado bajo la jurisdicción regulatoria de los reguladores laxos.

El traslape regulatorio no debería confundirse con un sistema mixto (de regulación más litigación). En el sistema mixto, los diferentes reguladores (la agencia administrativa y el tribunal) se apilan en orden vertical; el paralelo que se da en los negocios corresponde a la jerarquía en una firma. En un sistema de traslape regulatorio, dos o más agencias pueden encontrarse empoderadas para regular la misma actividad, y la esperanza es que la competencia entre ellas mantendrá a cada una pendiendo de un hilo.

[18] Véase Jean-Jacques Laffont y David Martimort, "Separation of Regulators against Collusive Behavior", 30 *RAND J. Econ*., 232 (1999).

[19] Jody Freeman y Jim Rossi, "Agency Coordination in Shared Regulatory Space", 125 *Harv. L. Rev.*, 1131 (2012).

[20] Teresa M. Schwartz, "Protecting Consumer Health and Safety: The Need for Coordinated Regulation among Federal Agencies", 43 *Geo. Wash. L. Rev.*, 1031 (1975).

XXIV.6. La captura del regulador

El término "captura del regulador" se refiere a la subversión que hacen de las agencias regulatorias las compañías a las que regulan. Debe distinguirse este fenómeno de la regulación cuyo propósito, asignado por el cuerpo legislativo que la promulga, consiste en servir a los intereses privados de las compañías reguladas, por ejemplo, protegiéndolas de la entrada de nuevos miembros al círculo empresarial. La captura implica conflicto, y la captura del regulador implica que las compañías reguladas hayan emprendido la guerra (por así decirlo) contra la agencia regulatoria y la hayan ganado, convirtiéndola en su vasallo. Los primeros en prestar una meticulosa atención a la captura regulatoria fueron los científicos políticos, sobre todo Marver Bernstein. En su libro *Regulating Business by Independent Commission* [*La regulación de la empresa a través de una comisión independiente*], Bernstein propuso la existencia de "un ciclo vital de comisiones reguladoras" que se dividía en cuatro fases: "gestación", "juventud", "madurez: el proceso de devitalización" y "vejez: debilitamiento y declinación".[21] La gestación es la promulgación del estatuto regulatorio tras "un periodo en el que lentamente se cierne una aflicción a causa de un problema".[22] Como este periodo está llamado a ser prolongado y a involucrar una fiera batalla entre intereses en pugna, el estatuto regulatorio bien puede resultar anticuado a la hora de ponerlo en funcionamiento.

De todas maneras, cuando finalmente es creada, la agencia regulatoria probablemente se incline a comenzar "su carrera administrativa siendo presidida por un espíritu agresivo, de cruzado".[23] Ésa corresponde, por supuesto, a la fase de la juventud. No dura mucho. Los grupos que batallaron *por* el programa regulatorio están cansados; el apoyo difuso del público no está a la altura de las empresas reguladas, que integran un grupo cohesionado cuyo objetivo es hacer a la dependencia su servidora; la agencia se empieza a embrollar en litigaciones con las empresas reguladas; la legislatura que creó a la agencia pierde el interés. A continuación, la agencia ingresa a su fase de madurez, en la que viene a formar parte del *statu quo* y se afana tanto en proteger su propia existencia y sus prerrogativas como en conciliar a la industria regulada: "Es improbable que la comisión, en este periodo, tenga la capacidad para extender la regulación más allá de los límites aceptables para los grupos regulados […] La comisión pierde su vitalidad […] Sus objetivos devienen un asunto de rutina y son una cosa ya aceptada […] Quizás el rasgo más marcado en el desarrollo de una comisión llegada a su madurez es el crecimiento de una pasividad que linda con la apatía. Es palpable un deseo de evitar conflictos a toda

[21] Marver H. Bernstein, *Regulating Business by Independent Commission,* cap. 3 (1995).
[22] *Ibid.,* en 74.
[23] *Ibid.,* en 80.

costa y de disfrutar de buenas relaciones con los grupos regulados".[24] El periodo de madurez culmina en "el acto mediante el que la comisión se rinde ante los grupos regulados […] Finalmente, la comisión se transforma en un cautivo de los grupos regulados".[25]

El análisis de Bernstein confunde el ciclo vital de las instituciones en general con el ciclo de vida distintivo (si es que lo hay) de las agencias regulatorias. Las instituciones, tanto las públicas como las privadas, tienden a osificarse, en particular si inicialmente alcanzan el éxito, debido a que a partir de entonces son renuentes al cambio; pueden preferir el riesgo de una lenta declinación al riesgo del súbito desastre que, por supuesto, pueden precipitar los cambios de rumbo. Por añadidura, conforme una institución crece de tamaño, como lo tendrá que hacer si se vuelve exitosa, el espíritu emprendedor cederá el paso a la administración, a la burocracia (lo que el gran sociólogo alemán Max Weber llamaba la rutinización del carisma), puesto que las prácticas burocráticas son indispensables para la administración de las organizaciones de tamaño considerable, sean públicas o privadas.

No obstante, la burocratización y la osificación, características clave del ciclo de la vida institucional, no tienen que implicar necesariamente una captura. La desviación de una agencia respecto de sus metas originales, aunque puede acompañar a la tendencia natural de una institución a burocratizarse, no constituye un factor inherente a ésta. Aun así, debe concederse crédito a científicos políticos como Marver Bernstein por haber estado interesados en las agencias regulatorias como organizaciones; en la era en que apareció su libro, los economistas no tendían a estarlo. Preocupados por las fallas registradas en el mercado, solían concebir la regulación como un método esencialmente libre de costo para rectificar esas fallas.

A la larga, un movimiento económico de signo conservador cuyo florecimiento se centró en la Universidad de Chicago comenzó a examinar las características sistémicas de la regulación que, en lugar de los errores de diseño o implementación a fin de cuentas evitables, podrían dar cuenta de las perversidades económicas detectadas de ordinario en la regulación tal como era aplicada. George Stigler propuso una teoría económica general, positiva, de la regulación,[26] la cual por cierto pasaba por alto la captura del regulador según el sentido que Bernstein le había asignado al término. De acuerdo con la teoría de Stigler, un ejemplo temprano de la teoría de la elección pública (véase antes XX.3), la regulación es un producto gobernado por las leyes de la oferta

[24] *Ibid.,* en 87-88.

[25] *Ibid.,* en 90.

[26] George J. Stigler, "The Theory of Economic Regulation", 3 *Bell J. Econ. & Mgmt. Science,* 3 (1971). Véase también Gary S. Becker, "A Theory of Competition among Pressure Groups for Political Influence", 98 *Q. J. Econ.,* 171 (1983); Sam Peltzman, "Toward a More General Theory of Regulation", 19 *J. Law & Econ.,* 211 (1976); Richard A. Posner, "Theories of Economic Regulation", 5 Bell *J. Econ. & Mgmt. Science,* 335 (1974).

y la demanda. El proveedor es la legislatura, o una agencia regulatoria en el caso de que la legislatura hubiera delegado la autoridad general a una agencia en lugar de especificar el programa regulatorio que la agencia había de administrar, y el demandador es algún grupo —presumiblemente, un grupo de consumidores o de empresarios— que anticipa recibir beneficios del programa regulatorio. Stigler creía que los demandadores provendrían usualmente de un grupo de empresarios y no de uno de consumidores, porque los beneficios de un programa regulatorio para un consumidor serían por lo general ligeros y los costos derivados de organizar a los consumidores en un grupo de presión efectivo serían elevados… y, por lo demás, ¿quién estaría interesado en sobrellevar esos costos? De esta manera, la principal competencia en pos de la regulación procedería de las industrias rivales o de segmentos de la industria.

En 1972, cuando se publicó el artículo de Stigler, los principales programas regulatorios eran, en efecto, específicamente intrínsecos a la industria. Y la mayoría de ellos eran de naturaleza cartelista, sobre todo la Comisión Interestatal de Comercio, la Junta de Aeronáutica Civil, la Junta Nacional de Relaciones Laborales (NLRB, la cual propició los monopolios laborales), la Comisión de Bolsa y Valores (que impuso vigorosamente un cártel de corredores de bolsa y de bolsas de valores), la Comisión Federal de Comunicaciones (que limitaba la competencia en las emisiones de la radio y la televisión y en la telefonía), la Comisión Federal de Comercio (que protegía a los minoristas frente a la competencia de las tiendas en cadena y las casas de descuento; protegía a los agentes corredores de la industria de la alimentación; protegía a los fabricantes de productos de piel contra los fabricantes de sustitutos de tela, y protegía a los fabricantes de bienes de consumo domésticos contra los productores extranjeros… "made in Japan"), las agencias reguladoras de bancos, las agencias de acreditación de los profesionistas e incluso los reguladores municipales de taxis. Con todo, el análisis de Stigler, así como virtualmente toda la literatura subsecuente dedicada a exponer y amplificar la teoría positiva de la regulación económica, eran de carácter estático. Un programa regulatorio era una mercancía que comparaba la industria regulada ("como regla, la regulación es adquirida por la industria y está diseñada y opera primariamente para su beneficio"[27]), si bien algunos análisis mostraron que la presión política ejercida por un grupo de consumidores había tenido influencia sobre el programa.[28] La posibilidad de que el programa pueda no haber tenido al inicio una naturaleza proteccionista, anticompetitiva, o de que se pueda haber alterado al paso del tiempo debido a factores organizacionales tales como la burocratización, no se incorporó a sus análisis.

El movimiento de desregulación alteró la teoría económica positiva de la regulación en un respecto. Marver Bernstein había tenido dificultades en su

[27] Stigler, nota 26, en 3.

[28] Véase, por ejemplo, Richard A. Posner, "Taxation by Regulation", 2 *Bell J. Econ. & Mgmt. Science,* 22 (1971).

teoría de "las edades del hombre" acerca de la regulación al abordar la era final, que debía haber sido la del debilitamiento y la muerte… pero es cierto también que escribió su obra antes de que las agencias comenzaran a pasar a mejor vida (con unas pocas excepciones particulares, como la de la Comisión Federal de Radio, que cedió el paso a la Comisión Federal de Comunicaciones, o la de algunas agencias de los tiempos de la Gran Depresión y la guerra). A resultas del movimiento de la desregulación, que tuvo sus inicios a fines del decenio de 1970 principalmente en reacción al patético funcionamiento de la economía en esa década, célebres agencias federales, como la Comisión Interestatal de Comercio y la Junta de Aeronáutica Civil, quedaron abolidas; otras fueron despojadas de gran parte de su autoridad, o bien se tornaron mucho menos proteccionistas (son notables ejemplos la Comisión de Bolsa y Valores, la Comisión Federal de Comercio y las agencias regulatorias de los bancos). No obstante, estos cambios no provenían del interior, como el envejecimiento biológico, que fue el modelo de captura del regulador aportado por Bernstein.

Mientras tanto, conforme la regulación tradicional intrínsecamente ligada a esta o aquella industria sucumbía a la desregulación, la regulación de las prácticas que trascendían el marco de industrias específicas, como la de la contaminación y otras externalidades ambientales, las condiciones laborales insalubres, los productos riesgosos o desfavorables a la salud, la explotación de los consumidores crédulos y la discriminación en el empleo, estaba aumentando. Estas agencias no han sucumbido a la desregulación ni han experimentado ciclos de vida a la usanza de Bernstein. Aun cuando se encuentran inmersas en una lucha a brazo partido con las empresas y otras instituciones dedicadas de lleno a las actividades objeto de su regulación, las agencias no exhiben ninguna tendencia general a sufrir la captura a manos de las compañías reguladas. Los programas regulatorios que han sobrevivido cuentan con un apoyo público que es considerable, y, en virtud de que estos programas no constituyen regulaciones amplias de las actividades competitivas de industrias específicas (esto es, regulaciones del precio, la entrada al mercado, la expansión y las obligaciones propias de la prestación de servicios), como ocurre en la regulación concerniente al transporte comunal o colectivo y los servicios públicos y que aportó el modelo dominante de las agencias regulatorias tradicionales estudiado por Bernstein y Stigler, resulta más dificultoso para las compañías sometidas a la regulación ejercer esfuerzos intensos y concentrados en la captura. La meta de los negocios al vérselas con las formas más novedosas de la regulación consiste justamente en lograr el debilitamiento de ésta… aunque cabe añadir que, usado en estos términos, el concepto es desorientador. La Ley de Comercio Interestatal, creada por la Comisión Interestatal de Comercio, fue promulgada en 1887 a instancias de los agricultores del Medio Oeste. La Comisión fue capturada primero por los ferrocarriles y fue utilizada para poner límites a la competencia entre ellos, y subsecuentemente por los camioneros, a fin de poner barreras a la competencia proveniente de los ferro-

carriles. Ésa es la auténtica narrativa de la captura: un fenómeno más interesante que los esfuerzos a menudo exitosos de las firmas reguladas en pos del debilitamiento de la regulación.

Lo que es verdad es que, debido a que los funcionarios del máximo nivel de las agencias regulatorias federales son designados por el presidente, los grupos de interés, ya sea que se trate de industrias, sindicatos o grupos de consumidores o ambientalistas, influyen en las agencias regulatorias. Uno puede concebir esta influencia como una especie de captura, aun cuando sería mejor elegir las palabras "infiltración" o "socavación".

XXIV.7. La estructura de la agencia administrativa

La agencia reguladora independiente ("independiente" en el sentido de que, a diferencia de agencias tales como la Agencia de Protección Ambiental de los Estados Unidos [EPA] que se hallan totalmente inscritas en la rama ejecutiva, sus miembros no rinden servicio a condición de complacer al presidente sino que pueden ser removidos sólo por una causa justificada), tal como la Comisión Federal de Comercio, la Comisión Federal de Comunicaciones o la Junta Nacional de Relaciones Laborales, es una combinación interesante de funciones legislativas, litigiosas y de cumplimiento. La justificación original era el deseo de librar al Congreso de algunas de las cargas de la legislación. El Congreso no podía afrontar eficientemente los problemas numerosos, muy técnicos, rápidamente cambiantes de una industria moderna compleja tal como la de los ferrocarriles; aun así, la función reguladora podría haberse delegado igualmente a los tribunales, cuyo papel tradicional es precisamente formular y aplicar reglas reguladoras de actividades que a menudo son complejas. Podría sostenerse que el método del caso concreto limita la eficacia de la elaboración de reglas judiciales, pero, dado que la mayoría de las agencias han dependido ampliamente de este método como su técnica legislativa, el argumento provee una base precaria para preferir las agencias sobre los tribunales. Ciertamente, las agencias se han mostrado más susceptibles a la influencia política que los tribunales federales. Su jurisdicción más especializada las somete a un escrutinio más estrecho por parte de los subcomités de asignaciones del Congreso, a través de los cuales se transmiten a la agencia las influencias políticas que se ejercen en el Congreso, y a una atención más estrecha por parte de las industrias reguladas por ella. La independencia política de las agencias administrativas es también menor que la de los jueces porque sus miembros tienen una duración limitada y la rotación es grande.

El verdadero propósito de la delegación a las dependencias no es el mejoramiento del funcionamiento técnico del proceso legislativo sino el aseguramiento de un cumplimiento más flexible de las políticas *no* motivadas por las metas de eficiencia en relación con lo que puede esperarse de los tribunales.

Recordando el análisis del poder judicial independiente del capítulo XX, podemos describir a las agencias administrativas como una forma del poder judicial "dependiente" destinada a promover la operación de la política de grupos de interés, antes que la eficiencia de la asignación, en calidad de delegada de la legislatura. Este enfoque sugiere tres proposiciones acerca de la regulación administrativa que tienen un apoyo empírico:

1. Las agencias administrativas se establecerán con mayor frecuencia cuando es grande la probabilidad de anular *de facto* la legislación por parte del aparato judicial (como ocurría, por ejemplo, durante el Nuevo Trato cuando los tribunales eran hostiles a la regulación económica federal).
2. Sin embargo, la legislatura conservará cierta revisión judicial de las determinaciones administrativas, si bien en forma atenuada, a fin de asegurarse de que la agencia, en su entusiasmo por servir a la legislatura actual, no se aleje demasiado de los términos del "trato" legislativo que se tradujo en la creación de la agencia en primer término (véase antes XX.5).
3. La adjudicación administrativa será menos congruente a lo largo del tiempo que la adjudicación judicial; una observación relacionada es que el precedente desempeñará en la decisión administrativa un papel menor que en la decisión judicial. Esto se infiere directamente del carácter relativamente dependiente del juicio administrativo.

Una innovación anunciada del proceso administrativo fue el relajamiento de la estructura de la agencia administrativa. Ésta podría expedir reglas, examinar casos, decidir casos, realizar estudios, dar asesoría y proponer legislación. Se pensaba que la combinación de funciones era una fuente de vigor. En la práctica, es probable que la combinación más importante haya sido la del procesamiento y la adjudicación.

La producción de decisiones administrativas involucra una secuencia de actividades: investigación, demanda, juicio, sentencia y apelación, de modo que la combinación de la expedición de la queja y la instrucción y la decisión de apelar ante el mismo organismo corresponde a la integración vertical en las empresas. Como sabemos, la justificación económica de la integración vertical en el contexto empresarial es que ella permite ahorros de costos mediante la sustitución del contrato por el mando como un método de coordinación de la producción (véase antes X.10). La esencia del problema de la coordinación en la fabricación de acero, por ejemplo, consiste en asegurarse de que el fabricante de hierro produzca y entregue al fabricante de acero la cantidad correcta de hierro, de la calidad correcta, en el momento correcto, donde "correcto" significa algo que está de acuerdo con las especificaciones del fabricante de acero. Si las compañías tienen propietarios diferentes, estas especificaciones

se escribirán en un contrato entre las compañías o serán objeto de una negociación *ad hoc*. Si las compañías tienen un mismo propietario, la coordinación se logrará de un modo diferente: los administradores de la empresa le dirán al fabricante de hierro cuánto debe producir, de cuál calidad, etcétera.

De modo similar, cuando el procesamiento no queda bajo el control de la agencia responsable del arbitraje de un conflicto, ésta tendrá que comunicarle al fiscal su "demanda" por quejas indirectamente, a través de reglas u opiniones que establezcan las doctrinas que acepta la agencia y los requerimientos probatorios que impone. Dicha comunicación tiene asignada la misma función que las especificaciones en el contrato de un comprador con un vendedor. El método alternativo de coordinación consiste en que la agencia controle la emisión de quejas. Luego puede vetar las quejas que tenga razones para dudar que pudiera sostener, y puede ordenar la emisión de quejas que juzgue meritorias e importantes. Éste puede no ser el método más eficiente de coordinación. La comunicación a través de reglas y decisiones puede ser más eficiente en virtud de que obliga a la agencia a articular políticas y prioridades; a planear por adelantado en vez de reaccionar a propuestas sobre quejas sostenidas por el personal de la agencia.

Asimismo, la coordinación es sólo uno de los elementos que deben considerarse en un análisis costo-beneficio de la elección entre la combinación y la separación de las funciones administrativas. Otro es el del sesgo, lo que en este contexto significa que la agencia pondera los costos de una decisión errónea para desechar una queja (*su* queja, cuando el procesamiento y el arbitraje de un conflicto se fusionan) en mayor medida que los costos de una decisión errónea para expedir una orden remedial contra el acusado. El caso extremo surge cuando la agencia otorga una ponderación nula a los costos del inocente castigado. Entonces adoptaría (si está en libertad de hacerlo) procedimientos que aseguraran la condena de todos los acusados. Estos procedimientos minimizarían los costos sociales impuestos cuando los violadores escapan al castigo y lo harían a un costo administrativo mínimo, ya que las cargas probatorias del enjuiciamiento serían ligeras y los derechos procedimentales del acusado serían nulos. Pero esta conclusión depende de si la agencia busca maximizar el número de condenas ponderadas por la severidad o bien la disuasión. Si se trata de esta última, la agencia será sensible a la reducción habida en la disuasión producida por la condena rotunda del inocente (ya que, situándose en el límite, si la probabilidad de ser castigado fuese idéntica para el inocente y para el culpable, el castigo no tendría ningún efecto disuasivo).

¿Deberíamos esperar un sesgo menor de una agencia donde estén separadas las funciones de procesamiento y adjudicación o arbitraje de un conflicto (como ocurre en la Junta Nacional de Relaciones Laborales, cuyo abogado general tiene un control completo del procesamiento y no es designado ni puede ser despedido por la Junta)? Una agencia que descarta muchas de sus propias

quejas se halla sujeta a la crítica de que, al traer a cuento casos no meritorios, exhibió a las claras su pobreza de juicio y dilapidó recursos de por sí escasos. No obstante, es improbable que la separación del procesamiento y la adjudicación tenga un efecto dramático en el sesgo. Una dependencia administrativa difiere crucialmente de un tribunal por cuanto está encargada del logro de cierto fin regulador sustantivo (la prevención de las prácticas laborales injustas, el engaño causado a los consumidores o el crecimiento del monopolio) y no sólo de la resolución de disputas. Y, en relación con las metas de la dependencia, la implementación de una orden punitiva o remedial tiene un peso inherentemente distinto del que tiene desestimar la queja. La implementación de la orden promueve la meta de la agencia; la desestimación no, en ningún sentido que sea probablemente percibido por quienes juzgan el trabajo de la agencia. Imaginemos que en un año la FTC rechazó todas las quejas que se le plantearon, quizá debido a que habían cambiado por completo las condiciones económicas que habían originado los estatutos aplicados por la agencia. Ésta estaría invitando a su liquidación por el Congreso. Tales perspectivas deben ser disuasivas. Los tribunales, con su cartera de casos diversificada, no tienen inhibiciones similares contra desestimar muchas o la mayoría de las quejas planteadas bajo un estatuto particular.

La inevitabilidad de que haya sesgo en la agencia es un factor que pesa en la renuencia de las legislaturas a concederle fuertes facultades remediales. El remedio administrativo normal es la orden de cesar y desistir, que en esencia es una prohibición, y la ausencia de otros remedios es una fuente de debilidad, como hemos visto (antes en XIV.2). Si las agencias pudieran imponer sanciones que infligieran costos muy grandes a los acusados, los costos sociales de la adjudicación sesgada por parte de la agencia serían mucho mayores que ahora.

XXIV.8. El comportamiento de las agencias administrativas

Una agencia administrativa puede ser ineficiente de dos maneras. Una es el hecho de tener metas (ya sea por su propia iniciativa o en cumplimiento del mandato judicial) contrarias a la promoción de la eficiencia (llamamos a esto "ineficiencia de la asignación"). La otra consiste en fracasar en la minimización de los costos de la persecución de sus metas, cualesquiera que ellas sean (la llamamos "ineficiencia interna"). Ya hemos examinado un fenómeno que según se cree soporta la acusación de ineficiencia interna: el hecho de que las agencias planteen muchos casos aparentemente triviales. Otros son que su personal con frecuencia recibe sueldos menores que sus contrapartes del sector privado y —lo que es una observación relacionada— que con frecuencia abandona la agencia en favor de empleos mejor pagados en la industria regulada por la agencia. Se obtienen de estos datos las inferencias de que el personal de la agencia es inferior, de que se ve sutilmente corrompido en el desem-

peño de sus obligaciones para con ella por la perspectiva de un empleo en la industria regulada, y de que su rotación de personal es excesiva. Éste es el problema bien conocido de la "puerta giratoria": al personal de las agencias lo pueden integrar aves de paso, cuyos incentivos como empleados se complican por sus ambiciones de hallar un puesto lucrativo en el sector privado.

Empero, son también posibles inferencias alternas, de carácter benigno, a partir de la frecuencia con que el personal de la agencia migra al sector privado. La primera es que dicho personal es contratado por la industria porque la capacitación especializada y la experiencia adquirida en el trabajo para la agencia aumentan su productividad en la industria regulada por comparación con otros empleos.[29] La segunda es que las agencias pueden atraer empleados competentes con sueldos menores que los que tendrían que pagar los empleadores privados precisamente porque la capacitación y la experiencia impartidos por la agencia aumentan el poder de ganancias de toda la vida del empleado: se trata de una inversión en su capital humano que el empleado paga bajo la forma de salarios temporalmente menores.[30] Lo que es verdad, no obstante, es que, en compensación por sus bajos salarios, los empleados de agencia, a semejanza de otros empleados gubernamentales, poseen invariablemente una gran seguridad laboral. Así, en realidad esos empleados no están tan mal remunerados como parece. Con todo, la seguridad laboral va aunada al efecto que tienen las mediciones ineficientes del desempeño al sofocar los incentivos para el rendimiento eficiente, reduce el control político sobre el personal de la agencia y, como veremos en la siguiente sección de este capítulo, vuelve a las agencias demasiado adversas al riesgo.

Así pues, haciendo un balance general, ¿es la "puerta giratoria" una cosa buena o mala? Es difícil decirlo. Por un lado, el empleado de la agencia que ambiciona pescar finalmente un empleo privado querrá demostrar que es vigoroso y agresivo, pues éstos serán activos importantes para su empleador privado. Por otro, este tipo de empleador puede dejar de contratar a los empleados de la agencia más agresivos en un esfuerzo por impedir la aplicación excesivamente rigurosa de las regulaciones.

Entre los mecanismos que impiden que una agencia y su personal desperdicien recursos se encuentra la ambición de los miembros de la agencia. Sus aspiraciones de ascensos o de empleos privados bien pagados se fortalecen si se ganan una reputación de eficiencia. Otro factor es el proceso de las asig-

[29] Ross D. Eckert, "The Life Cycle of Regulatory Commissioners", 24 *J. Law & Econ.,* 113 (1981).

[30] La agencia no resulta necesariamente perjudicada por el hecho de que los empleados de antaño con frecuencia representen a litigantes ante ella. Quienes solían ser sus empleados pueden tener habilidades especiales para asesorar a sus clientes con respecto a las verdaderas probabilidades de que la agencia prevalezca en el caso. Esto puede dar origen a una tasa de arreglo más elevada, lo que habrá de disminuir los costos de la agencia, así como a arreglos que se encuentren más cerca del precio óptimo que ella establezca para el arreglo.

naciones legislativas. La competencia de las agencias por las apropiaciones corresponde, aunque sólo aproximadamente, a la competencia de las empresas privadas por el capital. A la empresa ineficiente no le va bien en los mercados de capital; la agencia ineficiente pierde asignaciones ante su rival más eficiente. No obstante, dado que la producción de una agencia no se vende en un mercado, de modo que resulta difícil de evaluar, y debido a que los incentivos de los legisladores para poner riendas al gasto gubernamental son débiles (no logran quedarse con nada de los ahorros), la disciplina del proceso de las asignaciones es mucho más débil que la de los mercados de capital, aun cuando la agencia, a diferencia de una compañía de negocios, tiene que retornar a su mercado de capital cada año.

XXIV.9. Los costos de agencia en las agencias públicas

Sabemos que los costos de agencia —los costos derivados de la dificultad para alinear cabalmente los incentivos de un agente con los de su principal— son un problema incluso en las empresas privadas maximizadoras de la ganancia. El desempeño de la empresa misma es objetivamente medible por su rentabilidad, pero puede resultar difícil evaluar la contribución que haga un individuo como empleado a ese desempeño. Los costos de agencia revisten incluso más gravedad en las agencias públicas, en parte a causa del problema de la puerta giratoria discutido antes y en parte debido a la ausencia de una buena medición del desempeño de la agencia en su conjunto: una ausencia que distiende la presión sobre el director de la agencia para que minimice los costos. Un ejemplo perturbador de los efectos distorsionadores que conlleva la administración deficiente de una agencia administrativa es el de la Oficina de Patentes y Marcas Registradas de los Estados Unidos (USPTO o PTO). La práctica seguida en la PTO de exigir una declaración de razones por parte del examinador de patentes sólo cuando se rehúsa a aceptar una solicitud de patente, así como el hecho de que la PTO es financiada mediante cuotas aplicadas a patentes y para el mantenimiento de patentes, generan un fuerte sesgo institucional a favor de la autorización de dichas solicitudes. Esto contribuye a explicar el excesivo número de patentes concedidas, según se advirtió en el capítulo XI.

Otro estudio encuentra que las carreras subsecuentes de los fiscales federales de alto rango (los fiscales de los Estados Unidos) se ven influenciadas favorablemente por la duración de las sentencias obtenidas en su oficina y no por el número de condenas.[31] Aun así, presumiblemente, el Departamento de Justicia preferiría que los fiscales de los Estados Unidos ponderaran las

[31] Richard T. Boylan, "What Do Prosecutors Maximize? Evidence from the Careers of U.S. Attorneys", 7 *Am. Law & Econ. Rev.,* 379 (2005). Véase también Richard T. Boylan y Cheryl X. Long, "Salaries, Plea Rates, and the Career Objectives of Federal Prosecutors", 48 *J. Law & Econ.,* 627 (2005).

sentencias por las condenas. Éstos no proceden de esa forma porque unos pocos procesamientos exitosos que desembocan en sentencias prolongadas constituyen una sólida evidencia de la habilidad de dichos fiscales para cerciorarse de que una serie de casos pesados lleguen a completarse de manera exitosa, creando así una referencia valiosa para su reputación en la práctica privada.

Otro estudio más descubre que los asistentes de los fiscales federales de los Estados Unidos (quienes llevan a cabo efectivamente el procesamiento) tienen más probabilidades de llevar un caso a la instrucción que de negociar un acuerdo con el fiscal si se ubican en el área en la que los salarios legales son elevados.[32] Los salarios son uniformes en todos los distritos, y los empleos como asistentes de los fiscales federales de los Estados Unidos son profundamente codiciados, por lo que muchos abogados que deciden trabajar como asistentes de esa naturaleza en tales áreas tendrán más probabilidades de conducirse de esa manera como una forma de invertir en su capital humano, obteniendo una experiencia en el juicio de instrucción que les acarreará un mayor éxito en la práctica privada, que de buscar una carrera como fiscal. Empero, no se logran cumplir las metas del Departamento de Justicia atribuyendo a los juicios de instrucción un peso decisivo en la promoción de las carreras, a menos que se trate de mejorar el perfil del conjunto de solicitantes (aunque así se incrementa también la rotación).

Si, como en el caso del FBI, existen dos vías por donde transitan las carreras y una (la de las investigaciones penales) involucra mediciones objetivas del desempeño, mientras que la otra (la inteligencia de seguridad nacional) no, el personal más dotado de habilidades gravitará hacia la actividad objetivamente medible. Saben que les irá bien en lo económico si su desempeño se evalúa objetivamente, y también saben que, si su trabajo no se presta a las mediciones objetivas del desempeño, les podrán sacar ventaja en la competencia por la promoción colegas menos habilidosos que disponen de conexiones personales o políticas o bien son habilidosos en lo referente a lamer las botas de sus superiores.

Un costo de agencia endémico que conlleva el empleo público es la aversión al riesgo. Tal como sabemos, las agencias ofrecen una titularidad más segura que las compañías privadas, en parte porque se hallan menos sujetas a las vicisitudes del mercado y en parte porque cumplen la finalidad de proteger a su fuerza de trabajo contra la interferencia política. Esto conduce a la autoselección por parte del individuo adverso al riesgo para el empleo gubernamental, lo cual genera un costo de agencia debido a que no hay razón alguna para que una agencia pública sea adversa al riesgo en lugar de ser neutral.

[32] Boylan y Cheryl X. Long, "Salaries, Plea Rates, and the Career Objectives of Federal Prosecutors", 48 *J. Law & Econ.,* 627 (2005).

XXIV.10. La revisión judicial de las decisiones de las agencias

Tal como se mencionó anteriormente, las cortes poseen típicamente jurisdicción para revisar decisiones tomadas por las agencias administrativas. La revisión es de carácter deferente, en reconocimiento de que éstas se ocupan de temas técnicos que son de su competencia, pero, más realistamente, toman decisiones en materia de políticas bajo la guisa de amplias delegaciones legislativas, en vez de aplicar una ley a determinado hecho tal como lo haría una corte distrital.

La doctrina Chenery[33] prohíbe que un tribunal de revisión confirme la acción administrativa por una razón diferente de la de la agencia (de ordinario una justificación esgrimida por los abogados de la agencia al defender su decisión en el tribunal de revisión). Si se pensara que la agencia se ocupa simplemente de una búsqueda de la verdad, sería apropiado que el tribunal de revisión confirmara la decisión de la agencia por cualquier razón que tuviera sentido para el tribunal; presumiblemente, la agencia se adheriría a esa razón. Pero si son políticos los resortes reales de las decisiones de la agencia, resultará difícil para el tribunal pronosticar si ésta adoptaría la razón sugerida.

Una cuestión que se plantea con frecuencia consiste en saber si la revisión judicial de la acción de la agencia debiera realizarse inicialmente en un tribunal de distrito (de instrucción), con el derecho de apelar ante el tribunal de apelaciones, como en el caso de la revisión judicial de las negaciones de los beneficios por discapacidad a cargo de la Administración de Seguridad Social, o, como es más común, en el tribunal de apelación desde el principio, descartando al tribunal de distrito. Dos estratos de revisión aumentarán el número total de casos examinados en los tribunales, pero disminuirán el número de los casos presentados ante el tribunal de apelación. Supongamos que hay 100 decisiones administrativas, que se pide la revisión en 50 de ellas y que, si los tribunales de distrito tienen una jurisdicción de revisión inicial, 20% de sus decisiones de revisión se apelarán ante los tribunales de apelaciones. Entonces, en un sistema de dos estratos habrá un total de 60 casos, 50 en los tribunales de distrito y 10 en los tribunales de apelaciones, mientras que en el sistema de un solo estrato habrá sólo 50 casos, pero todos ellos en los tribunales de apelaciones. Si la revisión en los tribunales de apelaciones es más costosa para el sistema judicial, por las razones sugeridas en el capítulo anterior, es posible que el sistema de dos estratos sea más eficiente aunque el estrato adicional tenga escaso efecto en la disminución del número y, por ende, de los costos de los errores legales, siempre y cuando el sistema de dos estratos disminuya el número total de los procedimientos de revisión judicial, como puede hacerlo incrementando los costos de las partes.

[33] Véase SEC *vs.* Chenery Corp., 318 U.S., 80, 94 (1943).

Si suponemos que los jueces y los administradores son maximizadores racionales de la utilidad y que la confirmación interviene positivamente y la revocación negativamente en sus funciones de utilidad (véase antes XX.7), la relación existente entre las agencias y los tribunales de revisión tendrá un elemento competitivo. Por ello, cuando la Corte Suprema autorizó la revisión judicial de las reglas administrativas formales (las adoptadas después de la advertencia y su discusión pública) antes de que se aplicaran contra compañías particulares,[34] las agencias respondieron disminuyendo su utilización de reglas elaboradas tras una advertencia y su comentario, porque ello habría requerido ahora que la agencia creara un registro muy meticuloso a fin de soportar esa revisión, mientras que si la revisión se pospusiera hasta que se presentara un procedimiento de ejecución se podría desarrollar en ese momento un registro estrechamente vinculado con la controversia de que se tratara en ese procedimiento, y no antes. O supongamos que un tribunal de revisión puede elegir entre la revocación de la decisión de una agencia por razones de la interpretación estatutaria y porque ésta hubiera cometido un error procedimental o porque no hubiera expresado una justificación razonada de su acción. Si el tribunal escoge la primera vía, esto volverá más difícil que la agencia confirme su decisión tras una revisión, pero más fácil que la agencia persuada a la Corte Suprema para que revise y revoque la decisión del tribunal, ya que la Corte se interesa más por los problemas de interpretación estatutaria que por la corrección de errores en casos específicos.[35]

Lecturas recomendadas

Becker, Gary S. y George J. Stigler, "Law Enforcement, Malfeasance, and Compensation of Enforcers", 3 *J. Legal Stud.,* 1 (1974).

Bishop, William, "A Theory of Administrative Law", 19 *J. Legal Stud.,* 489 (1990).

Cohen, Mark A. y Paul H. Rubin, "Private Enforcement of Public Policy", 3 *Yale J. on Reg*., 167 (1985).

"Conference on the Economics and Politics of Administrative Law and Procedures", 8 *J. Law, Econ. & Organization,* 1 (1992).

DeMuth, Christopher C. y Douglas H. Ginsburg, "Rationalism in Regulation", 108 *Mich. L. Rev.,* 877 (2010).

Glaeser, Edward L., Daniel P. Kessler y Anne Morrison Piehl, "What Do Prosecutors Maximize? An Analysis of the Federalization of Drug Crimes", 2 *Am. Law & Econ. Rev.,* 259 (2000).

[34] Véase Abbott Laboratories *vs.* Gardner, 387 U.S., 136 (1967).

[35] Emerson H. Tiller y Pablo T. Spiller, "Strategic Instruments: Legal Structure and Political Games in Administrative Law", 15 *J. Law, Econ. & Organization,* 349 (1999).

Graetz, Michael J., Jennifer F. Reinganum y Louis L. Wilde, "The Tax Compliance Game: Toward an Interactive Theory of Law Enforcement", 2 *J. Law, Econ. & Organization,* 1 (1986).

Landes, William M. y Richard A. Posner, "The Private Enforcement of Law", 4 *J. Legal Stud.,* 1 (1975).

Posner, Richard A., "The Behavior of Administrative Agencies", 1 *J. Legal Stud.,* 305, 323-344 (1972).

——— "The Concept of Regulatory Capture: A Short, Inglorious History" (en *Preventing Capture: Special Interest Influence in Regulation, and How to Limit It* (Dan Carpenter y David Moss [comps.], 2013).

Research Handbook on Public Choice and Public Law, partes 1, 2 y 4 (Daniel A. Farber y Anne Joseph O'Connell [comps.], 2010).

Roberts, John, *The Modern Firm: Organizational Design for Performance and Growth* (2004).

Smith, Joseph L. y Emerson H. Tiller, "The Strategy of Judging: Evidence from Administrative Law", 31 *J. Legal Stud.,* 61 (2002).

Tiller, Emerson H. y Pablo T. Spiller, "Strategic Instruments: Legal Structure and Political Games in Administrative Law", 15 *J. Law, Econ. & Organization,* 349 (1999).

Problemas

1. ¿Cómo debieran decidir el Congreso y el Servicio de Impuestos Internos (IRS) cuánto gastar en la aplicación de las leyes fiscales federales? Supongamos que se estima que, el año pasado, los contribuyentes evadieron el pago de 50 000 millones de dólares en impuestos. ¿Debiera el gobierno gastar en consecuencia hasta 50 000 millones de dólares para recaudar estos impuestos? Supongamos que el irs no tiene idea acerca del monto de la evasión que hay en el total, pero sí sabe que por cada dólar adicional que gaste en el cumplimiento recaudará otros 300 millones de dólares. ¿Es ésta una razón suficiente para incrementar las asignaciones otorgadas a la dependencia? ¿Qué diremos si se aumentan tales asignaciones hasta el punto donde 1 dólar adicional de gasto en el cumplimiento genere sólo 65 centavos de dólar de recaudación adicional? ¿Es ésta una prueba concluyente de que el IRS está gastando demasiado ahora?

2. El IRS ofrece a los informantes una recompensa de hasta 10% de los impuestos no pagados que recaude como resultado de la ayuda del informante. ¿Es esto demasiado o muy poco? ¿Sería 100% una recompensa apropiada? ¿Podría defenderse una cifra mayor a 100%?

3. La cuestión de la demora en los procedimientos administrativos suscita una consideración del conflicto existente entre, por un lado, la idea de que las agencias administrativas se crean para servir al interés público, pero con frecuencia no lo hacen a causa de las malas políticas o del personal, y, por otro, la

idea de que las agencias están diseñadas en efecto para promover los intereses de grupos políticos eficaces, lo que hacen más o menos eficientemente. ¿Es la demora siempre ineficiente? ¿Puede imaginar algunos casos donde la demora ayude o perjudique a grupos de intereses particulares afectados por la acción administrativa? Véase antes XIII.3.

Séptima Parte

LA CONSTITUCIÓN Y EL SISTEMA FEDERAL

XXV. LA NATURALEZA Y LAS FUNCIONES DE LA CONSTITUCIÓN[1]

XXV.1. La economía del constitucionalismo: una introducción

La palabra "constitución" es rica en significados, pero en el contexto estadunidense, donde se hace hincapié naturalmente en la Constitución de los Estados Unidos, se emplea por lo general y se empleará aquí para denotar un tipo de legislación que: *1)* requiere una supermayoría o mayoría cualificada para ser cambiada y *2)* establece las facultades, las obligaciones y la estructura fundamentales del gobierno. Estos criterios están conectados. Entre más fundamentales sean los arreglos que una regla prescribe, menos deseable será probablemente el cambio frecuente.

Entre los temas en el análisis económico de las constituciones se incluyen:

1. La teoría económica del constitucionalismo: es decir, las propiedades y las consecuencias económicas del hecho de requerir una mayoría calificada para cambiar una ley.
2. La economía del diseño constitucional —de las reglas constitutivas de un sistema político— y, de este modo, de principios tales como la separación de poderes dentro del gobierno federal y la asignación de facultades entre el gobierno federal y los estados (federalismo).
3. Los efectos económicos (ampliamente definidos) de doctrinas constitucionales específicas tales como la regla de exclusión en cateos y confiscaciones o las limitaciones que la Corte Suprema, en nombre de la Constitución, ha establecido sobre los juicios por difamación contra la prensa y otros medios de comunicación.
4. La interpretación de las disposiciones constitucionales que parecen tener una lógica económica implícita, tales como la cláusula de la libertad de expresión cuando se concibe como una garantía de un mercado libre de ideas, así como también la cláusula de las confiscaciones, la cláusula del comercio y la cláusula de los contratos.
5. Las propuestas de reforma del derecho constitucional para volverlo una protección amplia de los mercados libres mediante la reinterpretación de disposiciones existentes mediante nuevas enmiendas, tales como la que requiere un presupuesto balanceado.

[1] Véase Geoffrey R. Stone *et al.*, *Constitutional Law* (6ª ed., 2009).

6. El problema del "doble rasero" constitucional: la paradoja de que la Corte Suprema proteja más la libertad personal que los derechos de propiedad, aunque la Constitución parezca otorgar mayor protección a los últimos que a la primera.
7. La relación existente, si es que la hay, entre la Constitución y el crecimiento económico.
8. Los principios de la interpretación constitucional, considerados desde el punto de vista de la teoría de principal-agente.

Este capítulo y los siguientes se ocupan de varios de estos temas.

La Constitución se estudia a menudo como una especie de contrato relacionado con otros no convencionales como el "contrato social". Desde un punto de vista económico resulta adecuada la analogía de las constituciones con los contratos. Al revés de una ley, un contrato puede modificarse sólo con el consentimiento de las partes (salvo en circunstancias excepcionales, como la del error mutuo); un cambio requiere el consentimiento unánime antes que la preferencia mayoritaria. Los costos de transacción harían de la unanimidad una condición inviable de la enmienda constitucional, pero puede considerarse que los requerimientos de una mayoría calificada para cambiar la Constitución de los Estados Unidos (una votación de dos tercios por ambas Cámaras del Congreso más la ratificación de las tres cuartas partes de los estados) trasladan a la enmienda constitucional hacia el extremo contractual del continuo ley-contrato.

Aun así, las constituciones, incluida la Constitución de los Estados Unidos, no son inmunes a las presiones de los grupos de interés que plagan al proceso legislativo ordinario. Si bien la dificultad de obtener la protección constitucional la vuelve más costosa, la dificultad correlativa de la revocación una vez que se obtiene aumenta su valor. Ese valor es el más grande si el grupo que ha buscado la protección constitucional espera perder poder político en el futuro, de modo que ese grupo no puede depender de su capacidad para evitar la legislación adversa mediante el proceso democrático normal,[2] o si (el caso más común) la ayuda legislativa se ve frenada por una disposición constitucional existente o por decisiones judiciales que interpretan tal disposición.

Sabemos que un contrato a largo plazo genera un problema de obsolescencia debido a las limitaciones de la presciencia humana. La solución reside en la flexibilidad de la interpretación, lo que requiere un lenguaje suficientemente general que dé paso a una interpretación que adapte el contrato a los cambios de las circunstancias no previstos al momento de su firma, así como principios de interpretación que autoricen a los tribunales a adoptar semejante flexibilidad en las lecturas. Los paralelos en la esfera constitucional son

[2] Donald J. Boudreaux y A. C. Pritchard, "Rewriting the Constitution: An Economic Analysis of the Constitutional Amendment Process", 62 *Fordham L. Rev.,* 111 (1993).

la bien conocida generalidad de varias de las disposiciones constitucionales más importantes, así como el principio de que la interpretación constitucional debe ser flexible antes que estricta. El otro lado de la moneda es el problema conocido de los costos de agencia. Entre más general sea la disposición que haya de interpretarse, y entre más laxa sea la interpretación permitida, más fácil será que el agente-intérprete (en este contexto, la Corte Suprema) se aleje de la ruta fijada por el principal (los artífices de la estructura de la Constitución).

Muchas disposiciones constitucionales se expresan en generalidades, pero algunas son muy específicas, como la disposición de que cada estado tiene derecho a dos senadores. Estas disposiciones tienen la naturaleza de las convenciones. Una convención reconocida establece que los conductores de automóviles deben manejar por la derecha. No es importante que sea la derecha antes que la izquierda, pero es importante que haya una adhesión congruente a cualquier regla arbitraria que se adopte. La disposición de los dos senadores es similar. Podrían ser igualmente uno, o tres, o cuatro. Dado que una convención es arbitraria, lo que quiere decir (sólo con una ligera exageración) que no varía respecto de su ambiente social (porque una convención contraria se ajustaría igualmente a ese ambiente), no hay necesidad de cambiarla en la medida que se modifica el ambiente. Por lo tanto, los costos de la especificidad son escasos. Pero los beneficios son grandes. Las disposiciones constitucionales precisas, como el requerimiento de los dos senadores, minimizan los costos de litigación (puesto que no hay nada de la incertidumbre que alimenta la litigación) y, lo que es más importante, disminuyen los recursos destinados al conflicto político al retirar algunas cuestiones de la arena de la política ordinaria. Si el número de los senadores se fijara por ley, habría esfuerzos frecuentes, algunos de ellos exitosos, para cambiar el número conforme, ora esta parte, ora aquélla, vieran una oportunidad para obtener o cimentar su control de la mayoría. Estos gastos, que tendrían un producto social escaso o nulo, se eliminan al volver prohibitivo el costo del cambio.

Cuando se contemplan desde un punto de vista económico, el diseño y la interpretación de las constituciones implican una tensión entre la eficiencia y la democracia. La eficiencia se maximiza por una constitución que confine al gobierno a tomar medidas para impedir externalidades negativas y alentar las externalidades positivas, y que insiste (en la medida de lo posible) en que, dentro de su esfera estrictamente circunscrita, el gobierno aplique políticas que minimicen los costos. Sin embargo, el espíritu gubernamental así concebido es el del gobierno limitado antes que el del gobierno democrático. No obstante, a menos que resulte fácil que el pueblo abandone la jurisdicción gubernamental si no está satisfecho con sus servicios, no hay nada que impida a los funcionarios públicos expandir su papel bastante más allá de lo prescrito por los principios del gobierno limitado. Aquí regresa el principio democrático. Añade la "voz" a la "salida" como una restricción al abuso del cargo gubernamental. Pero entre más congruentemente se aplique el principio democrá-

tico, mayor será el peligro de que una coalición asuma el control del gobierno y lo utilice para expropiar la riqueza de quienes no son miembros de la coalición, o de que los grupos de interés procuren conseguir una legislación que los enriquezca a expensas del público en conjunto.

Entre más ampliamente interprete la Corte Suprema las disposiciones constitucionales que limitan las facultades de los poderes elegidos, más moverá a la nación hacia el modelo del gobierno limitado; la Corte se vuelve un sustituto del pueblo en la limitación del gobierno. Entre más estrechamente interprete la Corte estas disposiciones, más moverá a la nación hacia el modelo de gobierno democrático… pero es un modelo congruente con la denigración y la explotación de quienes carecen de poder político.

Los artífices de la Constitución estaban muy conscientes de los peligros de la democracia. En la Carta Magna original, sólo los miembros de la Cámara de Representantes eran elegidos directamente por el pueblo; y el electorado estaba restringido. La mala distribución de las legislaturas estatales generaba sustanciales disparidades en el poder de votación efectivo entre los individuos, hasta que se declaró inconstitucional a principios del decenio de 1960. El argumento en favor del sufragio universal es triple: *1)* la exclusión de cualquier grupo provocará ciertos esfuerzos (los que son socialmente costosos y de todos modos no producen ninguna ganancia social) para redistribuir la riqueza de sus integrantes entre otros grupos electoralmente poderosos; *2)* las elecciones generan información, valiosa para los gobernantes, acerca de las preferencias y las aversiones (en una era de encuestas científicas, este argumento es menos importante que antes), y *3)* entre mayor sea el electorado, más difícil será la formación de coaliciones para fines redistributivos (esto se relaciona con el primer argumento). Sólo el segundo argumento explica por qué quienes tienen el voto habrían de tomar la decisión de diluir su propio poder agrandando la franquicia, a menos que fuese probable que el grupo excluido recurriera a la fuerza si no se le otorgara pacíficamente una porción del poder político. Pero quizá se expanda la franquicia siempre que el grupo dominante en la comunidad crea que quienes están excluidos ahora prefieren a sus candidatos frente a los de grupos rivales. Una vez obtenida la franquicia, un grupo podrá utilizar su poder de votación para oponerse a una nueva exclusión. Así que la expansión del sufragio tiene un trinquete, de modo que raras veces se revierte.

XXV.2. La separación de poderes

La Constitución trata de impedir la centralización del poder gubernamental dividiendo el poder político entre los estados por una parte y el gobierno federal por la otra, y separando los poderes del gobierno federal en tres ramas independientes: ejecutivo, legislativo y judicial. Hay también división dentro de cada rama. El poder legislativo se divide entre la Cámara de Representantes,

el Senado y el presidente con su poder de veto; el poder ejecutivo y el judicial dependen del legislativo para sus asignaciones; el poder judicial depende del ejecutivo para la ejecución de los fallos, y así sucesivamente.

En términos económicos, el propósito principal de la separación de poderes es la monopolización del poder coercitivo del Estado, que es una forma de monopolio potencialmente mucho más costosa que cualquiera otra de las discutidas en este libro. La observación formulada en el capítulo x, en el sentido de que los costos de la conspiración para los conspiradores aumentan con el número de aquellos cuyo consentimiento es necesario para volver eficaz la conspiración, se aplica a las conspiraciones políticas tanto como a las comerciales. La aplicación de la fuerza coercitiva del gobierno según el sistema considerado en nuestra Constitución requiere la concurrencia del poder legislativo, para que promulgue la medida coercitiva; la del poder ejecutivo, para realizar una acción de cumplimiento (de ahí la importancia de un monopolio público de la aplicación de la ley federal en algunas áreas: faculta a la rama ejecutiva para desempeñar una función de contrapeso sobre el Congreso al no hacer cumplir al pie de la letra las promulgaciones de éste), y la del poder judicial, para aplicar una sanción por el incumplimiento por las acciones legales del ejecutivo o el legislativo. Tres no es un gran número de conspiradores, pero sólo uno de los tres —el ejecutivo— es unitario aun en principio.

Un aspecto importante de la separación de poderes es el requerimiento de que algunos ciudadanos particulares seleccionados al azar (los jurados) concurran en la imposición de una sanción penal o en el otorgamiento del pago de daños civiles, a menos que el demandado renuncie a su derecho al juicio por jurados. Y compartir el poder del gobierno nacional con numerosos cuasisoberanos (los estados) complica enormemente los esfuerzos que se hacen para crear un monopolio del poder político.

El hecho de que los artífices de la Constitución separaran deliberadamente los poderes principales del gobierno plantea una interrogante acerca de la legitimidad del proceso administrativo discutido en el capítulo precedente. Se defiende el proceso porque fortalece la eficiencia, pero el fortalecimiento se obtiene (presuntamente) combinando el poder ejecutivo, el legislativo y el judicial en un solo organismo. Éste parece el tipo de eficiencia que los artífices de la Constitución consideraban ineficiente en general a causa del peligro que creaba de una centralización excesiva del poder político. Es cierto que el derecho de las partes a un procedimiento de agencia, a fin de obtener la revisión judicial de su decisión, limita el poder administrativo. (Por la misma razón, las medidas que se han adoptado, sobre todo en la ley del procedimiento administrativo, para dotar de importancia al derecho de la revisión judicial de la acción de la agencia disminuyen la eficiencia del proceso administrativo.) Pero la revisión judicial no provee ninguna restricción al ejercicio del poder legislativo por una dependencia administrativa a través de la elaboración de reglas bajo una delegación amplia del poder del Congreso:

una práctica diseñada para evadir un sistema deliberadamente complicado de promulgación legislativa.

Las agencias "independientes" (véase antes XXIV.7) son una cuarta rama de gobierno, y, aunque a menudo sus críticos las tachan de inconstitucionales porque la Constitución generó sólo tres ramas o poderes (ejecutivo, legislativo y judicial), en la medida en que son independientes (sólo parcialmente, por supuesto) de las otras ramas realmente aumentan el grado de separación de los poderes; pero las agencias administrativas que no son independientes, que se encuentran incorporadas plenamente a la rama ejecutiva, disuelven la separación de poderes dentro del dominio de las agencias al empoderar al Ejecutivo para que ejercite facultades judiciales y legislativas tanto como ejecutivas. Por ejemplo, al expandir la deferencia judicial a la elaboración de leyes administrativas en su decisión Chevron,[3] la Corte Suprema debilitó sin querer la separación de poderes.

La separación de poderes reduce la eficiencia al incrementar los costos de transacción gubernamentales pero la incrementa al permitir una explotación más completa de la división del trabajo. Si los jueces federales operaran al gusto de los políticos, sería difícil atraer personas capaces a los puestos de juzgados federales; la ley sería aún menos estable de lo que es, porque cambiaría con los vientos políticos, y al Congreso le resultaría difícil alcanzar tratos duraderos con grupos de interés (véase antes XX.5). De igual modo, el Congreso es demasiado poco manejable por sus dimensiones para poder hacer la ejecución de las políticas, mientras que la rama ejecutiva, si se la autorizara a legislar, carecería de la capacidad de deliberación y la sensibilidad política pragmática del Congreso, y éstos son recursos valiosos en la promulgación de la legislación. Los jueces que legislan incorporando en la ley sus preferencias políticas, o que usurpan las prerrogativas del ejecutivo, como cuando se dice a las agencias de la rama ejecutiva dónde asignar los recursos de la aplicación, olvidan las ventajas de la división del trabajo.

Como hemos comenzado a advertir, el término "separación de poderes" es engañoso porque sugiere que cada rama del gobierno realiza facultades no ejercidas por las otras ramas, mientras que de hecho existe un considerable traslape, en particular entre el Congreso y el presidente. Los principios de la organización mercantil pueden aportar cierta guía sobre cómo dividir responsabilidades entre las ramas legislativa y ejecutiva. Si concebimos al presidente como el director ejecutivo de la rama ejecutiva y al Congreso como la junta directiva llamada a representar a la población estadunidense, siendo sus miembros concebidos como los "accionistas" del gobierno, entonces la función propia del Congreso en relación con las agencias de la rama ejecutiva debería ser la de supervisar antes que la de prescribir en detalle sus estructuras administrativas. Y, del mismo modo, podemos concebir al Congreso en ana-

[3] Chevron U.S.A., Inc. *vs.* Natural Resources Defense Council, 467 U.S., 837 (1984).

logía con una junta de administradores fiduciarios de una universidad y al presidente en analogía con el rector de una universidad, o trazar otras analogías con organizaciones no gubernamentales.

Éstas son, en efecto, analogías pertinentes. A diferencia de una junta directiva o de una junta de administradores fiduciarios de cara a un presidente ejecutivo o un rector de universidad, el Congreso no nombra al presidente y no puede removerlo salvo con gran dificultad a través del proceso de destitución. Y una junta corporativa u otra junta privada no se asigna dinero para cubrir gastos operativos: éstos son cubiertos gracias a los rendimientos de la corporación, o al ingreso procedente de la matrícula, las donaciones y las ganancias derivadas de las inversiones de una universidad, en tanto que las agencias gubernamentales, debido a que generan un ingreso escaso o nulo, requieren de subvenciones anuales (a veces en periodos más cortos) de la legislatura. Con todo, estas diferencias no afectan el punto estructural básico de que, en una asignación racional de responsabilidades para la administración de una empresa, los detalles, incluidos los detalles organizacionales, de la instrumentación de una política son apropiadamente asignados al ejecutivo en vez de a un comité de empleados de medio tiempo, la totalidad o la mayoría de quienes son extraños en lo que toca a la empresa.

Asimismo, los miembros del comité están obligados a carecer de conocimientos especializados. Ésta es una consecuencia de las limitaciones de tiempo, así como de conocimientos y experiencia, que se les imponen. A semejanza de los miembros de las juntas directivas en una corporación o de las juntas de administradores fiduciarios de una universidad, ningún miembro del Congreso está dedicado de tiempo completo a supervisar a las agencias ejecutivas particulares. Es más: debido a las demandas que pesan sobre su tiempo para rendir servicio al electorado, recaudar fondos, organizar campañas, votar sobre la gama completa de legislaciones consideradas por el organismo al que pertenecen y participar en múltiples comités y subcomités, y asimismo debido a los límites fijados al periodo para formar parte de esos comités, pocos miembros del Congreso disponen del tiempo libre para forjarse un conocimiento básico de las funciones cumplidas por esta o aquella rama ejecutiva. Su situación, en efecto, es peor que la de los miembros de juntas corporativas (se asemeja más a la de los miembros de las juntas universitarias) porque la mayoría de los miembros de una junta corporativa son hombres de negocios y con frecuencia están involucrados en un negocio similar al de la corporación de cuyo consejo forman parte. Incluso pueden ser funcionarios de la corporación, aun cuando, como sabemos, prevalece una tendencia en contra de mantener a directores en puestos internos porque poseen información confidencial y esto pone en riesgo su función de supervisión (otro ejemplo de los beneficios que acarrea la separación de responsabilidades ejecutivas y de supervisión).

XXV.3. La protección de los derechos

La Constitución confiere derechos en contra del gobierno, tanto federal como estatal, además de conferir derechos al gobierno federal. Un derecho constitucional se parece mucho, analíticamente, a los derechos de propiedad discutidos en el capítulo III. Es un derecho de no interferencia, aunque sólo por el gobierno, que se hace valer a través de los tribunales. Tal derecho tiene una forma negativa, como ocurre con un derecho de propiedad. ¿Debieran reconocerse también derechos constitucionales positivos, tales como el de una educación pública y un nivel razonable de protección policiaca y de bomberos, mediante la interpretación de las disposiciones constitucionales redactadas en términos generales tales como las cláusulas del debido proceso y de la protección igualitaria? Probablemente no, porque los derechos constitucionales implican costos de información judicial mucho mayores que los derechos negativos. La ejecución de un derecho positivo requiere de gastos gubernamentales de índole sustancial, en detrimento de otros programas gubernamentales; y el esfuerzo de balancear diferentes reclamaciones sobre el fisco federal supera la capacidad informativa del poder judicial. Es cierto, los derechos negativos también requieren a veces gastos gubernamentales para ser enteramente eficaces. El derecho a la asesoría legal y el derecho al aborto son dos ejemplos. Para una persona rica, ambos derechos son negativos: son derechos contra la interferencia gubernamental en la contratación particular de un abogado y de un médico, respectivamente. Pero, para una persona pobre, estos derechos tienen que ser tanto positivos como negativos, porque una persona pobre no puede contratar a un abogado ni comprar un aborto sin la asistencia pública, y deberá depender de las vicisitudes de la caridad.

Muchos de los derechos protegidos por la Constitución se relacionan con el propósito antimonopólico de la separación de poderes; limitan el poder del gobierno para intimidar a sus oponentes. Incluso la cláusula de la justa compensación de la Quinta Enmienda puede interpretarse bajo esta luz: impide que el gobierno confisque en su totalidad la propiedad de sus opositores políticos, al conceder a los dueños de ésta derechos a la compensación en caso de que el gobierno proceda de ese modo. Mientras que la protección de expresar jucios políticos en la Primera Enmienda se relaciona aún más claramente con la práctica de poner cotos al poder del gobierno sobre sus opositores políticos, no puede decirse lo mismo de la libertad de publicar libros sexualmente francos que no tengan un contenido político. Esa libertad es un fin —un aspecto de la utilidad de las personas—, antes que un medio para prevenir la monopolización del poder político, y la protección constitucional que disfrutan podría reflejar la influencia de grupos de interés.[4] Esta sugerencia tiene

[4] Véanse algunas pruebas en Robert A. McGuire y Robert L. Ohsfeldt, "Self-Interest, Agency

implicaciones para varias cuestiones constitucionales, dos de éstas las discutiremos aquí, mientras que otra es el tema de la sección siguiente.

1. Se sugiere a veces que la protección de la libertad de expresión y de prensa debiera limitarse a la expresión política, según la teoría de que la única función de la Primera Enmienda que se relaciona con el propósito central de la Constitución es la protección del proceso electoral por el que se selecciona a los miembros del Congreso y al presidente. La concepción mucho más amplia del alcance de la Primera Enmienda, adoptada por la Corte Suprema en los últimos decenios (como se discute en el capítulo XXVIII), es congruente con la concepción de la Primera Enmienda también como una forma de la legislación protectora a favor de un grupo de interés integrado por intelectuales, editores, periodistas, panfletistas y otros que obtienen ingreso pecuniario y no pecuniario de la publicación y la propaganda.
2. A veces se pregunta si no es una perversión del principio constitucional el hecho de invocar una disposición constitucional en nombre de una mayoría, no de un grupo minoritario, como cuando se cuestiona la discriminación inversa (véase adelante XXVIII.5), o la discriminación contra las mujeres, o cuando se cuestionan los programas de mala distribución legislativa adoptados mediante un referendo popular. Pero un grupo grande será a menudo menos eficaz que uno pequeño, desde el punto de vista político, debido a los costos mayores de la acción colectiva para aquél (véase antes XX.3), de modo que el grupo grande necesitará en efecto más, no menos, de la protección constitucional contra la legislatura.

Otra perspectiva económica sobre los derechos constitucionales los considera designados para impedir las formas particularmente severas y costosas de la redistribución de la riqueza, como cuando se confisca la riqueza de una persona sin compensación, o se la convierte en esclavo, o se le impide practicar su religión. Al colocar tales exacciones fuera del poder de la legislatura, la Constitución disminuye los riesgos creados por el poder político sobre la distribución de la riqueza (definida en términos amplios), y por la misma razón disminuye los gastos de autoprotección de los grupos amenazados. Esto provee una posible justificación para la jurisprudencia de la Corte Suprema sobre los "derechos fundamentales", sobre todo cuando el grupo protegido no tiene las características de un grupo de interés efectivo.[5]

Theory, and Political Voting Behavior: The Ratification of the United States Constitution", 79 *Am. Econ. Rev.,* 219 (1989); *cf.* Fred. S. McChesney, "A Positive Regulatory Theory of the First Amendment", 20 *Conn. L. Rev.,* 355 (1988).

[5] Véase Lynn A. Stout, "Strict Scrutiny and Social Choice: An Economic Inquiry into Fundamental Rights and Suspect Classifications", 80 *Geo. L. J.,* 1787 (1992).

XXV.4. El alcance de los derechos constitucionales: el caso del terrorismo

¿Cuán amplia o cuán estrecha debe ser la definición que se dé a un derecho constitucional? Desde una perspectiva económica, y utilizando a guisa de ejemplo los derechos constitucionales de las personas de quienes se sospecha que son terroristas, el alcance óptimo de semejante derecho es el punto en el que una ligera expansión dentro de ese margen sustraería más de un valor (la seguridad pública en el caso del terrorismo) del que añadiría a otro (la libertad personal en el caso del terrorismo). Ese punto habrá de cambiar (ha cambiado) conforme las amenazas a la libertad y la seguridad sufren altibajos. No obstante, el equilibrio depende no solamente de las ponderaciones asignadas a cada valor sino también del efecto causado en cada una de las medidas desafiadas en cuestión. Una reducción a gran escala de la seguridad puede dominar a una reducción de pocas proporciones en la libertad incluso si ésta es concebida como algo mucho más valioso que la seguridad. Supongamos que la libertad vale 1 000 y la seguridad sólo 100; sin embargo, una reducción de 20% en la seguridad por invalidar una medida defensiva (como la de mantener detenidos e incomunicados a los sospechosos de terrorismo) producirá más daño ($100 \times 0.20 = 20$) que una reducción de 1% en la libertad como resultado de defender la medida ($1\,000 \times 0.01 = 10$).

Por supuesto, esto constituye un análisis marginal, y en términos legales de naturaleza convencional corresponde a la toma de decisiones guiada por un estándar y su rechazo de una toma de decisiones guiada por una regla. Los libertarios civiles desean generalmente que la constitucionalidad de las medidas de seguridad sea determinada por reglas y no por estándares, por ejemplo, la regla de que la expresión o discurso político se suprima únicamente si conlleva una incitación a cometer un delito. Les preocupa que los estándares hayan de dar a los jueces demasiada libertad para que sea posible deferir casos a medidas restrictivas impuestas por el Congreso o por el presidente. Empero, quieren que la *regla* esté basada en la libertad y que sólo la *excepción* lo esté en la seguridad. Una excepción en seguridad pública a una regla libertaria en materia civil proporciona menos protección para la seguridad nacional que un estándar en el que ni la libertad ni la seguridad tienen prioridad.

El equilibrio preciso de valores en pugna requiere que los tribunales presten seriamente atención a los riesgos en lugar de estar siempre insistiendo en la certidumbre. Sería un error creer que debe rechazarse una medida en particular porque no sabemos si habrá de ocurrir otro ataque terrorista contra los Estados Unidos a menos que se adopte esa medida y, así, no podemos saber con certeza si realmente ésta habrá de incrementar la seguridad. La mayoría de las medidas de seguridad van encaminadas a reducir los riesgos en vez de ir dirigidas a eliminar las certezas, pero ello no hace a las medidas injustificadas. Que uno no pueda saber si el acto de interrogar a un terrorista deter-

minado prevendrá un ataque no priva de valor a los interrogatorios hechos a los terroristas. Este punto queda escondido cuando existe un cierto perjuicio causado a un lado de la balanza y hay un beneficio incierto en el otro. Un terrorista sometido a la tortura equivale a incurrir en un perjuicio que conlleva certeza. Pero que la tortura acarree un beneficio para la seguridad nacional constituye sólo una probabilidad cuando da comienzo la tortura. La tortura puede resultar inefectiva ya sea porque la persona torturada no posee información valiosa que transmitir o porque tiene capacidad para resistir la tortura; en cualquiera de los casos, por añadidura, puede hacer que sus interrogadores emprendan la búsqueda de una aguja en un pajar al decirles mentiras.

Aunque no debe hacerse caso omiso de los beneficios probabilísticos, éstos deben ser descontados por la probabilidad de que realmente hayan de materializarse. El valor esperado de 10% de oportunidades de ganar una apuesta de 100 dólares es de sólo 10 dólares. Los libertarios civiles hacen hincapié en los efectos probabilísticos en la misma medida en que lo hacen los expertos en seguridad nacional. Están menos preocupados por el impacto de una disminución de las libertades civiles sobre aquellos individuos cuyas libertades se han reducido (digamos, una persona investigada a causa de sus creencias políticas) que por el impacto causado sobre los rasgos esenciales del paisaje social en su conjunto, tales como el vigor a largo plazo del discurso político. Si el peligro que el terrorismo plantea a la seguridad nacional es materia de especulación, también lo es el riesgo que las medidas antiterroristas plantean a la libertad.

Los costos de información son importantes a la hora de elaborar los derechos constitucionales. Es posible que los jueces no sepan lo suficiente acerca de las consecuencias de algunas medidas particulares para proteger la seguridad nacional como para poder mantener un equilibrio apropiado al respecto. Los miembros del poder judicial, a diferencia de los de las ramas ejecutiva y legislativa, no cuentan con una maquinaria para llevar a cabo un estudio sistemático y amplio de un problema. Tienen un personal exiguo. La judicatura tiene que aguardar a que llegue un caso para comenzar su indagación sobre los hechos y las ramificaciones políticas, y la presión que supone su carga de trabajo le exige decidir sobre el caso sin poder darse tiempo para estudiar los antecedentes, las circunstancias y las consecuencias probables. Si el caso involucra un tema como el de los contratos o los accidentes o los delitos ordinarios en los que los tribunales tienen una larga experiencia, sus limitaciones epistemológicas no serán severas. No obstante, si, como en los casos que implican el terrorismo, el tema es nuevo, la corte no tendrá el tiempo o los recursos para examinarlo con toda la pausa del mundo. Y, por otra parte, nuestros jueces, incluidos los jueces de la Corte Suprema, son generalistas. Los casos que implican un asunto de seguridad nacional integran sólo una porción pequeña de sus expedientes. No les sobra tiempo para dedicarlo a esos menesteres. Tanto el Congreso como la rama ejecutiva tienen bastante más información sobre seguridad nacional que el poder judicial.

XXV.5. Revisión de la racionalidad

Hay un antiguo debate sobre si la Corte Suprema debiera emplear las cláusulas del debido proceso y la protección igualitaria para derogar la legislación estatal que, aun sin restringir un derecho constitucional específico tal como el de la libertad de expresión, resulta poco razonable cuando se juzga por algún criterio general del bienestar social o del interés público. El análisis económico sugiere que la afirmación de tal poder por la Corte Suprema cambiaría fundamentalmente la naturaleza del proceso político democrático. Vimos en el capítulo XX que no hay ninguna presunción de que la legislación se promulgue a fin de promover el interés público. Un producto característico de un proceso legislativo democrático —tal vez de cualquier proceso— es el de la redistribución injusta de la riqueza en favor de grupos de interés políticamente eficaces. El interés público no intervendrá necesariamente en el proceso, excepto como una fachada para aumentar los costos de información de la oposición o la derogación. El magistrado Hugo Black, antiguo senador que tenía una comprensión clara del papel de los grupos de interés en el proceso legislativo,[6] rechazaba también —y como un corolario necesario— la noción de que la legislación que no invadiera ninguno de los intereses específicamente designados para su protección en la Constitución debiera ser todavía juzgada razonable para que se la promulgara.[7]

Se ha sugerido que la Corte Suprema debiera invalidar, bajo la cláusula de la protección igualitaria, cualquier ley donde el método escogido para el logro del propósito declarado de la ley sea poco apropiado.[8] Una ley que obligue a los peluqueros a tener una capacitación médica, ostensiblemente para mejorar la salud pública, sería inválida por esta razón. Pero si el Estado revisara la ley para declarar que su propósito era el aumento del ingreso de los peluqueros al limitar la entrada a la profesión, se apoyaría a la ley revisada. Esta propuesta disminuiría los costos de información de la oposición a la legislación de intereses especiales; pero su eficacia para disminuir la cantidad de ella dependería de la importancia de tales costos en la promulgación de esa legislación y (lo que es una observación relacionada) de la capacidad de los legisladores para elaborar preámbulos plausibles. Aun si la propuesta fuese eficaz, sería todavía susceptible a la objeción de que supone, sin ninguna base en el lenguaje o la historia de las disposiciones constitucionales relevantes, que la Constitución está diseñada para corregir los problemas del "gorroneo"

[6] Véase, por ejemplo, Eastern Railroad Presidents Conference *vs.* Noerr Motor Freight, Inc., 365 U.S., 127 (1961).

[7] Ferguson *vs.* Skrupa, 372 U.S., 726, 729 (1963).

[8] Gerald Gunther, "The Supreme Court 1971 Term-Foreword: In Search of Evolving Doctrine on a Changing Court: A Model for a Newer Equal Protection", 86 *Harv. L. Rev.,* 1, 20-21, 23 (1972).

en el proceso legislativo aun cuando tales problemas no conduzcan a una invasión de ninguno de los derechos protegidos ni minen la separación de poderes.

En realidad, la palabra "racionalidad" está mal empleada, por lo menos desde el punto de vista del economista, para describir lo que hacen los tribunales en estos casos. La legislación que invalidan, o que se les pide invalidar, es racional, sólo que no responde al interés público. El propósito principal del concepto de la protección igualitaria no es imprimir más racionalidad a la legislación; como vimos antes, en VIII.9, es el de aumentar los costos políticos de la legislación opresora impidiendo que la legislatura se concentre en un grupo estrechamente definido que podría carecer de poder político. A fin de tener efecto sobre los miembros del grupo, la legislatura debe provocar la indignación de otros grupos que ocupan la misma posición desde el punto de vista del interés público; el aumento del número de víctimas significa el aumento de la oposición.

Adviértase la implicación de que debiera prohibírseles a los legisladores crear excepciones mediante el instrumento que consiste en "eximir" a los participantes existentes en la actividad contra la que va dirigida la legislación. La cláusula de exención por derechos adquiridos le permite a una legislatura precisar la carga que les impone el estatuto a los grupos desfavorecidos, en tanto que los favorecidos quedan exceptuados de la aplicación del estatuto, esto es, bajo el resguardo de la cláusula de exención por derechos adquiridos.

Lecturas recomendadas

Aranson, Peter H., "Models of Judicial Choice as Allocation and Distribution in Constitutional Law", 1990 *BYU. L. Rev.,* 745, 807-820 (1990).

Boudreaux, Donald J. y A. C. Pritchard, "Rewriting the Constitution: An Economic Analysis of the Constitutional Amendment Process", 62 *Fordham L. Rev.,* 111 (1993).

Brennan, Geoffrey y James M. Buchanan, *The Reason of Rules: Constitutional Political Economy* (1985).

Cooter, Robert D., *The Strategic Constitution* (2000).

McGuire, Robert A. y Robert L. Ohsfeldt, "Self-Interest, Agency Theory, and Political Voting Behavior: The Ratification of the United States Constitution", 79 *Am. Econ. Rev.,* 219 (1989).

Posner, Richard A., "The Constitution as an Economic Document", 56 *Geo. Wash. L. Rev.,* 4 (1987).

———, *Not a Suicide Pact: The U.S. Constitution in a Time of National Emergency,* cap. 1 (2006).

Silver, Morris, "Economic Theory of the Constitutional Separation of Powers", 29 *Pub. Choice,* 95 (1977).

Voigt, Stefan, "Positive Constitutional Economics: A Survey", 90 *Pub. Choice,* 11 (1997).
Wittman, Donald, "Why Democracies Produce Efficient Results", 97 *J. Pol. Econ.,* 1395 (1989).

Problema

El presidente puede elegir siempre entre presentar un tratado internacional al Senado para someterlo a su aprobación (lo que requiere un voto de dos terceras partes), a ambas cámaras para su aprobación como un acuerdo legislativo-ejecutivo (lo que requiere un voto mayoritario en cada cámara), o no presentarlo a ninguna de las cámaras, en cuyo caso es un acuerdo puramente ejecutivo, aunque se reconoce todavía como vinculante en el derecho internacional. ¿Qué determinará la elección del presidente? Véase John K. Setear, "The President's Rational Choice of a Preratification Pathway: Article II Treaty, Congressional-Executive Agreement, or Executive Agreement?", 31 *J. Legal Stud.,* 5 (2002).

XXVI. EL DEBIDO PROCESO ECONÓMICO

XXVI.1. La libertad contractual como un principio constitucional

De acuerdo con lo establecido en la Quinta y la Decimocuarta Enmiendas (la primera de las cuales rige sobre el gobierno federal, la segunda sobre el gobierno estatal y el local) en sus cláusulas acerca del debido proceso, el gobierno no privará a una persona de su vida, su libertad o su propiedad sin que medie el debido proceso legal. Durante medio siglo, periodo que concluyó a finales de la década de 1930, la Corte Suprema invalidó, si bien de manera intermitente,[1] una serie de leyes estatales y federales que regulaban la actividad económica,[2] recurriendo al argumento de que las restricciones a la libertad de contrato, si han de ser congruentes con el debido proceso, deben ser justificables sobre la base de la teoría económica clásica tal como la expuso, por ejemplo, John Stuart Mill. En *Sobre la libertad,* Mill sostuvo que el gobierno tenía derecho a interferir con la libertad de conducta de una persona solamente si su comportamiento le causaba daño a alguien más. En consecuencia, la Corte Suprema invalidó las leyes sobre el máximo de horas, medida que alcanzó resonancia en Lochner *vs*. Nueva York:[3] un caso que se hizo célebre a causa de la disensión expresada por el magistrado Holmes y que dio su nombre a la "era de Lochner", tal como se llegaron a conocer los tiempos en que la idea sobre la libertad contractual cobró auge en la Corte. No obstante, ésta apoyó las leyes antimonopólicas y aquellas que sujetaban a los monopolistas a los controles de la tasa máxima (unos estatutos que, según se creía comúnmente, tenían como mira la eficiencia).

Aunque consideradas desde hace largo tiempo simplemente como distorsiones grotescas del principio constitucional, las decisiones de la era de Lochner han atraído recientemente a algunos partidarios acérrimos como parte de un resurgimiento creciente del interés por los principios económicos clásicos.[4] Y, aunque hay dificultades graves para conciliar su posición con la filosofía de la autorrestricción judicial o con la teoría (y la realidad) del gobierno

[1] David P. Currie, *The Constitution in the Supreme Court: The Second Century: 1888-1986*, 7-50 (1990).

[2] Véase Robert G. McCloskey, "Economic Due Process and the Supreme Court: An Exhumation and Reburial", 1962 *S. Ct. Rev.,* 34.

[3] 198 U.S., 45 (1905).

[4] Véanse, por ejemplo, Michael Conant, "Antimonopoly Tradition under the Ninth and Fourteenth Amendments: Slaughter-House Cases Re-Examined", 31 *Emory L. J.,* 785 (1982); Richard A. Epstein, *Takings: Private Property and the Power of Eminent Domain* (1985); Epstein, "Taxation, Regulation, and Confiscation", 20 *Osgoode Hall L. J.,* 433 (1982); Epstein, "Toward a Revi-

basada en el grupo de interés, lo mismo puede decirse del hincapié moderno en el derecho constitucional de las libertades personales. Son superficiales los argumentos legales a favor de otorgar una protección mayor a las libertades personales que a las económicas. Así, aunque se dice que no había ninguna fuente para una doctrina de la libertad contractual en el texto o en la historia de las disposiciones constitucionales relevantes, la misma crítica puede formularse (y se ha formulado) contra las decisiones de la Corte en una gran diversidad de otras áreas constitucionales. Se dice también que las cuestiones económicas son más difíciles de decidir por los tribunales que las cuestiones referentes a los derechos de los acusados de delitos, los disidentes políticos o los miembros de minorías raciales… aunque en efecto se sabe menos acerca de esas cuestiones que acerca de los problemas económicos convencionales. Se dice que los derechos económicos son menos importantes que otros derechos; aunque esto fuese cierto (una cuestión que examinaremos más adelante en XXIX.6), no significa esto que la Corte no debiera protegerlos o cuando mucho debiera conferirles una protección mínima. Se dice que el error de la Corte en los casos de libertad contractual era que estaba en desacuerdo con la opinión pública dominante. Pero esto fue cierto sólo hacia el final de la época, y fue por ello que la época terminó entonces. Además, la crítica puede convertirse fácilmente en un elogio para la Corte por su firmeza frente a una opinión popular contraria. Se dice también que las víctimas de los controles económicos son empresarios bien pertrechados para protegerse sin el auxilio de la Corte, al revés de lo que ocurre con las minorías indefensas involucradas de ordinario en un caso constitucional no económico. Pero, como veremos enseguida, la carga de la legislación económica cuestionada durante la época de la libertad contractual era soportada a menudo por grupos políticamente desorganizados tales como el de los consumidores. Tampoco es cierto que las minorías raciales y religiosas no puedan competir eficazmente en la arena política.

Casi a la manera de quien repara en un detalle —pero un detalle importante en vista del tema de este libro—, se sostiene la creencia común de que las decisiones de la libertad contractual reflejaban un entendimiento escaso de la economía. Una crítica temprana basada en esta decisión se encuentra en la opinión disidente del magistrado Brandeis en New State Ice Co. *vs.* Liebmann.[5] Este caso implicaba la constitucionalidad de una ley estatal que obligaba a quien quisiera fabricar y vender hielo a obtener un certificado de conveniencia y necesidad públicas y establecía que se negaría un certificado si el servicio existente era adecuado. New State, que tenía tal certificado, trató de lograr que se excluyera a Liebmann, quien no lo tenía, del territorio comercial de New State. La defensa de Liebmann consistió en que esa ley era in-

talization of the Contract Clause", 51 *U. Chi. L. Rev.,* 703 (1984); Bernard H. Siegan, *Economic Liberties and the Constitution* (1980).

[5] 285 U.S., 262 (1932).

constitucional. La Corte invalidó la ley por razones que aceptarían la mayoría de los economistas:

> Formulado sucintamente, una empresa privada trata aquí de impedir que un competidor entre al negocio de la fabricación y venta de hielo […] No se nos plantea ahora la existencia de ninguna regulación estatal que proteja al público consumidor respecto de las condiciones de la fabricación y la distribución, o para asegurar la pureza de los productos o impedir la extorsión. El control afirmado aquí no protege contra el monopolio, sino que tiende a promoverlo. No se trata de alentar la competencia sino de impedirla; no se trata de regular el negocio sino de impedir que algunas personas se involucren en él […] No es el caso de un monopolio natural, o de una empresa que por su misma naturaleza dependa del otorgamiento de privilegios públicos. Evidentemente, el requerimiento particular que se nos plantea no se impuso para impedir un monopolio práctico del negocio, ya que su tendencia es precisamente lo contrario.

La Corte asimiló la disposición de la certificación al intento de un zapatero, de acuerdo con la autoridad estatal, "por impedir que otro zapatero fabricara o vendiera zapatos porque los zapateros que ya se dedican a esa ocupación pueden fabricar y vender todos los zapatos que se necesiten".

El argumento económico del magistrado Brandeis se inicia con la proposición de que el negocio del hielo podría ser "uno que se presta peculiarmente para el monopolio"; "el negocio se realiza en plantas locales con un mercado estrechamente limitado en su área" en virtud del peso y el carácter perecedero del producto. Pero el hecho de que una empresa tenga sólo un área de mercado local no impide la competencia. Además, la opinión de Brandeis revela que antes de la promulgación de la ley combatida había competencia en el negocio del hielo en muchas localidades del estado. Sostiene Brandeis que, "aun en esas localidades, los precios del hielo eran ordinariamente uniformes", pero dado que, como destaca en otra parte de su opinión, el producto es uniforme, sería de esperarse que los vendedores competitivos cobraran el mismo precio.

La prueba del monopolio natural es el tamaño del mercado en relación con las condiciones de la oferta (véase antes XIII.1). Pero la opinión de Brandeis es oscura y contradictoria en lo tocante a la cuestión crítica de las condiciones de la oferta en el negocio del hielo. Subraya "la relativa facilidad y baratura con la que puede construirse una planta de hielo", así como el hecho de que el aumento de la producción de hielo no ha "tenido el efecto de aumentar grandemente el tamaño de las plantas en el negocio del hielo". Estas observaciones sugieren que la disposición competitiva de hielo no es más costosa que la monopólica. Pero también señala que "las plantas de hielo tienen una capacidad determinada, así como costos fijos y de operación inflexibles", lo que implica la producción en condiciones de costo medio decreciente.

No hay duda de que el verdadero propósito de la ley era el de promover la cartelización de la industria del hielo de Oklahoma. Como destaca curiosamente el propio Brandeis:

> Las revistas profesionales y los reportes de las reuniones de la asociación de fabricantes de hielo ponen ampliamente de manifiesto la hostilidad de la industria hacia tal competencia, y sus esfuerzos incesantes —a través de asociaciones profesionales, acuerdos informales, combinación de sistemas de reparto y, en particular, mediante la consolidación de plantas— para proteger los mercados y los precios contra la competencia de cualquier clase.

Observa igualmente que "toda la industria del hielo de Oklahoma ha aceptado la ley y el estatus que crea". ¡Naturalmente!

Al considerar el caso como uno en el que se enfrentaban los derechos económicos de Liebmann con los intereses de los pobres de Oklahoma que no podían comprar refrigeradores, el magistrado Brandeis lo hizo al revés. El derecho que habría reivindicado era el interés de New State Ice, y otras compañías de hielo establecidas, de verse libres de la competencia. La gente efectivamente perjudicada por la ley eran los pobres, quienes se veían obligados a pagar más por el hielo; los ricos, como señalaba Brandeis, tenían mayores probabilidades de contar con refrigeradores.

Si el negocio del hielo hubiese sido un monopolio natural, la posición de Brandeis podría ser económicamente defendible, ya que el esfuerzo de un monopolista natural por maximizar sus ganancias mediante el establecimiento de un precio monopólico podría conducir a una duplicación dispendiosa de las instalaciones al atraer nuevas entradas (véase antes XIII.1). Pero no sólo es falsa la premisa, sino que en la parte final de la opinión de Brandeis se descubre que el lenguaje de monopolio natural de la primera parte es una construcción artificial y que él estaba dispuesto a aceptar la proposición generalizante de que la competencia ruinosa es un fenómeno común en los mercados económicos y fue un factor importante tras la Gran Depresión del decenio de 1930. El caso se decidió en 1931, y, aunque la ley de Oklahoma era anterior a la Depresión, Brandeis discute extensamente, y con evidente aprobación, la proposición de que la filosofía incorporada en la limitación de Oklahoma a la entrada al negocio del hielo podría ser un remedio de aplicación general a la crisis económica en curso.

Está desacreditada la idea de que la Gran Depresión se generó en el exceso de competencia y podría curarse reduciendo la competencia. Por supuesto, cuando disminuyó la demanda durante la Depresión, se volvió temporalmente excesiva gran parte de la capacidad industrial existente, que se enfocaba en abastecer a una demanda incrementada. Pero limitar la competencia no habría aumentado el poder de compra y, por ende, la demanda; sólo habría afectado la eficiencia de la actividad económica a su nivel reducido. A pesar

de todo, el remedio del cártel para las depresiones se intentó en las leyes de principios del New Deal, como la Ley Nacional de Recuperación Industrial, que autorizaba a las industrias a fijar precios mínimos. Pronto se movió hacia el otro lado el péndulo de la opinión pública, sosteniéndose que el monopolio, antes que la competencia, había sido responsable de la Depresión o por lo menos de su gravedad. Este argumento no es más convincente que el anterior. Un monopolista reducirá su precio más lentamente que una empresa competitiva cuando disminuya el volumen de la demanda.[6] Pero, por otra parte, los precios deben bajar más deprisa en otros mercados, porque el precio relativamente elevado que el consumidor continúa pagando por el producto monopolizado reduce la cantidad de dinero que puede gastar en productos producidos competitivamente, lo que acelera la disminución de la demanda y (si suponemos costos marginales crecientes[7]) el precio de tales productos. Es improbable que sea significativo el efecto del monopolio sobre el nivel general del precio y, por lo tanto, sobre el poder de compra agregado del consumidor.

Algunas de las leyes confirmadas por la Corte Suprema en el periodo en el que se guiaba por las nociones de la libertad contractual fueron también intentos de suprimir la competencia bajo el disfraz de promover el bienestar general. Por ejemplo, en Muller *vs.* Oregon[8] la Corte confirmó una ley estatal que fijaba una jornada de trabajo máxima de 10 horas para las mujeres empleadas en las lavanderías. Es probable que la ley tuviese un efecto escaso, a menos que el Estado tuviera también una ley de salario mínimo y los salarios de las mujeres empleadas en las lavanderías no fuesen significativamente mayores que el mínimo. Obligado a reducir la jornada de trabajo, el empleador se compensaría reduciendo el salario diario. Si se impidiera esto último al empleador, trataría a la ley como un incremento del costo de su mano de obra (obtiene menos producción por el mismo salario), y, de acuerdo con un análisis que ahora es familiar, se adaptaría comprando una cantidad de mano de obra menor, elevando los precios, o haciendo ambas cosas. La disminución del empleo perjudicaría a los trabajadores despedidos que no tuvieran oportunidades de otro empleo igualmente bueno; el aumento de los precios perjudicaría a los consumidores, y al reducir su producción haría que el empresario disminuyera más aún sus insumos de mano de obra.[9]

[6] Véase antes X.1; Richard D. Reimer, "A Comment on Oligopoly Pricing Practices and Economic Theory", 38 *J. Bus*., 210 (1965).

[7] Si son constantes, el precio permanecerá invariable y, por consiguiente, la cantidad declinará.

[8] 208 U.S., 412 (1908).

[9] Hay pruebas de que tales leyes redujeron el empleo femenino, en particular el empleo de mujeres inmigrantes, dado que los trabajadores marginales son lastimados por una ley que incrementa el precio efectivo que debe pagar el empleador por la mano de obra. Elisabeth M. Landes, "The Effect of State Maximum Hours Laws on the Employment of Women in 1920", 88 *J. Pol. Econ.*, 476 (1980). *Cf.* antes XII.8.

West Coast Hotel Co. *vs.* Parrish[10] es uno de los casos en los que la Corte Suprema, en el famoso "cambio a tiempo que salvó a nueve", abandonó el intento por defender la libertad de contrato frente al estado regulatorio. Se discutía una ley estatal que fijaba un salario mínimo para las mujeres. La Corte Suprema confirmó la ley, destacando (sin citar prueba alguna) que "la negación de un salario de subsistencia" a las mujeres arrojaría "sobre la comunidad una carga directa para su sostenimiento", porque "lo que estos trabajadores perdieran en salarios tendría que ser pagado por los contribuyentes". Esto está al revés. Un salario mínimo, que deja sin empleo a los trabajadores cuyo producto marginal se encuentre por debajo de ese salario (véase antes XII.8), aumenta la demanda de beneficencia pública.

Después del repudio de las decisiones de la era de Lochner, la Corte ha confirmado repetidamente algunas leyes destinadas a promover el monopolio, como una ley estatal que, en aras de la salud pública, prohibió que las ópticas remplazaran los aros de los lentes sin una receta firmada por un optometrista o un oftalmólogo.[11] La ley podría no haber tenido otro propósito que el de aumentar los ingresos de los optometristas y los oftalmólogos a expensas de las ópticas y los consumidores.

XXVI.2. El resurgimiento del debido proceso económico: los pobres como una clase constitucionalmente protegida

En Griffin *vs.* Illinois,[12] la Corte Suprema confirmó que el requerimiento de que el acusado en un juicio penal comprara la transcripción de su juicio como una condición para obtener la revisión en apelación de algunas cuestiones dependientes de tal transcripción discriminaba contra quienes no pudieran pagarla. La Corte podría haber logrado el mismo resultado sosteniendo que el debido proceso en los juicios penales requiere que el acusado tenga un derecho de apelación y que ese derecho implica una oportunidad, si es necesario a expensas del Estado, para que se revise de manera *significativa* la apelación, para lo cual se requiere una transcripción del juicio. No obstante, si, como la opinión de Griffin parece sostener, un acusado pobre tiene derecho a una transcripción gratuita simplemente para neutralizar una ventaja que de otro modo tendría el acusado rico sobre él, se abren nuevas concepciones de la obligación constitucional. Si el Estado debe neutralizar las ventajas conferidas por la posesión de dinero en esta área, ¿por qué no en otras? En efecto, se ha sostenido que Griffin fue el primero de una serie de casos que pueden explicarse mejor con la teoría de que el gobierno está obligado, por un principio

[10] 300 U.S., 379 (1937).

[11] Williamson *vs.* Lee Optical Co., 348 U.S., 483 (1955); véase en general Pontarelli Limousine, Inc. *vs.* City of Chicago, 929 F.2d, 339 (7th Cir., 1991).

[12] 351 U.S., 12 (1956).

constitucional, a satisfacer las "necesidades mínimas justas", independientemente de la capacidad para pagar.[13] Se dice que otra decisión de la serie fue la de Harper, que invalidó el impuesto de capitación sobre el votante.[14]

Sin embargo, estas decisiones pueden explicarse por razones más precisas. Griffin puede asimilarse a los casos de garantía de asesoría legal para los acusados indigentes. Un individuo que (quizá sin culpa de su parte) es pobre tiene todavía un derecho de propiedad valioso: su libertad de acción. Si el gobierno puede quitárselo arbitrariamente —que es una forma de caracterizar un procedimiento penal donde sólo el gobierno tiene asesoría legal o donde el acusado no puede apelar eficazmente su condena—, disminuirá el valor de ese derecho de propiedad.[15] En cuanto a Harper, dado que la razón de beneficio-costo de la votación para el individuo es tan baja de todos modos, aun un impuesto de capitación moderado podría impedir que muchos pobres votaran. Tendría entonces el mismo efecto práctico que la prohibición de votar para los pobres, lo que sería incongruente con los principios del sufragio universal que pueden tener una justificación económica, como vimos en el capítulo anterior.

El desarrollo posterior de las implicaciones de Griffin es la invalidación por varios tribunales estatales, principiando por la Corte Suprema de California en Serrano *vs.* Priest,[16] de la práctica de financiar la educación pública con los impuestos prediales locales. Esta práctica hace de la cantidad de dinero por alumno gastada para la educación pública en cada distrito escolar una función, en parte, del valor de los inmuebles en el distrito. Se argumenta que, dado que los distritos más ricos contienen más inmuebles valiosos por regla general, los niños de padres más ricos reciben una educación más cara que los niños de padres más pobres, lo que podría ser ineficiente al igual que poco equitativo, ya que el propósito de la educación pública es el de invertir en capital humano y la inversión óptima depende de la inteligencia del niño y no de su riqueza (véase antes V.4). Con todo, la correlación existente entre la base del impuesto predial y el ingreso de las familias que utilizan las escuelas públicas no es siempre necesariamente positiva. En la ciudad de Nueva York, por ejemplo, hay muchos ricos y muchos inmuebles valiosos. Pero, dado que los ricos no envían a sus hijos a las escuelas públicas y gran parte del impuesto predial se recauda sobre la propiedad comercial antes que la residencial, el financiamiento de las escuelas públicas con los impuestos prediales tiene el efecto de redistribuir el ingreso en favor de los pobres. Si los gastos

[13] Frank I. Michelman, "On Protecting the Poor Through the Fourteenth Amendment", 83 *Harv. L. Rev.,* 7 (1969), criticado en Ralph K. Winter, Jr., "Poverty, Economic Equality, and the Equal Protection Clause", 1972 *S. Ct. Rev.,* 41.

[14] Harper *vs.* Virginia Bd. of Elections, 383 U.S., 663 (1966).

[15] En XIV.3 se sugirió antes otra justificación económica del derecho a la asesoría legal.

[16] 487 P.2d, 1241 (Cal., 1971). La Corte Suprema de los Estados Unidos se negó a seguir a Serrano. San Antonio Independent School Dist. *vs.* Rodríguez, 411 U.S., 1 (1973). Por lo que toca al análisis económico, véase Robert P. Inman y Daniel L. Rubinfeld, "The Judicial Pursuit of Local Fiscal Equity", 92 *Harv. L. Rev.,* 1662 (1979).

públicos por alumno se igualaran en todo el estado, la ciudad de Nueva York sería clasificada como un distrito rico y se reducirían sus gastos para escuelas públicas. Los beneficiarios principales de la igualación serían los habitantes rurales.

XXVI.3. Los derechos del debido proceso de los consumidores y los empleados públicos

El nuevo despertar judicial del interés por los derechos económicos se ejemplifica también por dos líneas de casos que a primera vista parecen convencionales en su insistencia en las salvaguardias puramente procedimentales (véase antes XXII.1). En Fuentes *vs.* Shevin,[17] la Corte Suprema invalidó ciertas leyes estatales que autorizaban, con el requisito de otorgar una fianza, la reivindicación (embargo), sin notificación o audiencia previa, de un inmueble que reclame como legalmente suyo la persona que busque tal reivindicación. Los demandantes eran vendedores de contratos de venta en abonos; los demandados eran compradores que presuntamente habían dejado de pagar. La Corte razonó que, dado que los compradores tenían un derecho de posesión sobre los bienes vendidos en los términos del contrato, los bienes eran de su propiedad en el sentido de la cláusula del debido proceso de la Decimocuarta Enmienda, de los que no podrían ser privados por la acción estatal sin notificación y una oportunidad de ser oídos previamente.

La premisa básica de la decisión de la Corte —que el "derecho de una persona para disfrutar lo suyo, libre de la interferencia gubernamental", es un derecho digno de la atención judicial—[18] congenia con el análisis económico, pero la conclusión de la Corte no. La Corte no cuestionó la reivindicación como un remedio apropiado para la falta de pago en un contrato de venta en abonos (véase antes XIV.14). Se trataba de saber cómo podría impedirse mejor que el remedio fuera invocado como un método de acoso cuando haya efectivamente una falta de pago. Desde el punto de vista de la economía, el método mejor consiste en crear un desincentivo para reivindicar sin razón. Las leyes cuestionadas hacían eso al requerir que los vendedores otorgaran fianzas. Aunque el procedimiento de reivindicación, como el procedimiento empleado para obtener una orden de restricción preliminar, es sumario inicialmente, no se puede llegar a una decisión final sin una audiencia. Si en la audiencia se revela que el inmueble fue reivindicado indebidamente, el vendedor deberá devolverlo y pagar los daños del comprador, si los hay; el vendedor debe pagar también de su bolsa el costo de la fianza. Por lo tanto, no tiene nada que ganar, y sí puede perder dinero si invoca el remedio sin razón. Dado que el

[17] 407 U.S., 67 (1972).

[18] *Ibid.,* en 81. Un panegírico más extravagante aún de los derechos de propiedad aparece en Lynch *vs.* Household Finance Corp., 405 U.S., 538, 552 (1972).

requerimiento de que el vendedor otorgue una fianza debiera desalentar las invocaciones frívolas del remedio y proteger al comprador cuando se cometan errores, los compradores tienen poco que ganar del hecho de tener también un derecho a una audiencia previa.

La Corte señaló el incentivo económico del vendedor para evitar invocar el remedio sin razón, pero sostuvo que tal incentivo no es un sustituto del fallo de un funcionario neutral. La preferencia de la autoridad sobre el interés propio como un regulador de la conducta humana resulta curiosa para un economista, quien también juzgaría extraño que la Corte debiera trazar una distinción tan nítida entre la cosa particular reivindicada y su equivalente económico. El requerimiento de que el vendedor otorgue una fianza y justifique la reivindicación en una audiencia subsecuente protegía el valor de los intereses de los compradores, aunque no la posesión continua del bien específico. En cuanto a los bienes de consumo comunes implicados en estos casos, el bien y su valor de mercado son intercambiables. Fuentes tuvo el efecto de aumentar el costo del contrato de ventas en abonos… lo que es un beneficio dudoso para los consumidores.

La segunda línea de casos, que también procede de la premisa inobjetable de que la Constitución prohíbe que el gobierno prive de propiedad a una persona sin el debido proceso legal, sostiene que los maestros y otros empleados públicos con contratos de permanencia —contratos que prohíben el despido arbitrario— tienen un derecho de propiedad en sus empleos, de modo que si son despedidos sin una audiencia se habrán violado sus derechos constitucionales. El uso del término "derecho de propiedad" en estos casos es cuestionable. La permanencia es un derecho contractual, no un derecho de propiedad. Lo opuesto a la permanencia no es un contrato de empleo sino el empleo a voluntad: el empleo sin salvaguardas contractualmente garantizadas y legalmente ejecutables contra la terminación (véase antes XII.5). En todo caso, es extraño que los tribunales piensen que estos empleados necesitan las protecciones constitucionales federales. Si se despide a un empleado público que tiene permanencia, podrá demandar al Estado en los términos de la ley contractual estatal. Y si el Estado no otorga remedios legales decentes a un empleado público despedido por incumplimiento del término de permanencia del contrato de empleo, los salarios de los empleados públicos serán mayores de lo que serían si existiesen tales remedios. Si los remedios costaran menos que la prima salarial que debe pagarse para compensar su ausencia, el Estado tendrá un incentivo para crear tales remedios. Si costaran más, el propio empleado preferiría (*ex ante*) no tenerlos. Y, dado que el derecho que se hace cumplir en un juicio según los términos de la Quinta o de la Decimocuarta Enmienda para hacer respetar un contrato de permanencia es un producto de ese contrato, y que los remedios contra el incumplimiento pueden considerarse como un término implicado del contrato, no está claro de cuál derecho de propiedad se ha privado al empleado si recibe los remedios a los que tiene derecho

contractualmente, aunque tales remedios no incluyan una audiencia que satisfaga los estándares federales del debido proceso.

Lecturas recomendadas

Epstein, Richard A., *Takings: Private Property and the Power of Eminent Domain* (1985).

Hovenkamp, Herbert, "The Political Economy of Substantive Due Process", 40 *Stan. L. Rev.,* 379 (1988).

Pritchard, A. C., "Government Promises and Due Process: An Economic Analysis of the 'New Property'", 77 *Va. L. Rev.,* 1053 (1991).

Scott, Robert A., "Constitutional Regulation of Provisional Creditor Remedies: The Cost of Procedural Due Process", 61 *Va. L. Rev.,* 807 (1975).

Siegan, Bernard H., *Economic Liberties and the Constitution* (2ª ed., 2005).

Winter, Ralph K., Jr., "Poverty, Economic Equality, and the Equal Protection Clause", 1972 *S. Ct. Rev.,* 41.

Problemas

1. ¿Utiliza la Corte Suprema las mismas premisas económicas en la decisión de casos antimonopólicos que en las decisiones de los casos modernos de derechos económicos, tales como Fuentes *vs.* Shevin?

2. Supongamos que se ha demostrado convincentemente que la Ley Federal del Salario Mínimo, a los niveles actuales, disminuye significativamente las oportunidades de empleo de los pobres (véase antes XII.8). ¿Debiera invalidarse en consecuencia tal ley como una violación de la protección igual?

XXVII. LA ECONOMÍA DEL FEDERALISMO[1]

XXVII.1. La asignación de responsabilidades entre el gobierno federal y los estados

El gobierno de los Estados Unidos se ha sobreimpuesto a los gobiernos de 50 estados. La razón de ello es que la Constitución era un pacto entre las 13 colonias británicas, las cuales, habiéndose sacudido el yugo británico y adquirido el estatus de estados soberanos independientes, decidieron formar una nación, pero no sin retener algunos poderes de autoridad suprema.[2] Así pues, se hicieron llamar "estados". La Constitución también autorizó que se admitieran estados adicionales en el futuro.

Un efecto potencial del acto de reconocer el estatus cuasisoberano de los estados es el de limitar el poder federal, pero ello depende de que entre éstos prevalezcan intereses comunes; si entre unos y otros hay intereses en pugna, cada uno buscará tener el respaldo federal, y el resultado consistirá en amplificar en vez de limitar el sistema federal.[3]

Una pregunta recurrente es la de si esta o aquella responsabilidad gubernamental recaen en el poder federal o en el estatal. Considerada desde el ángulo económico, la elección implica un trueque de sacrificios y ganancias entre tres factores:

1. El poder monopólico del gobierno. El gobierno federal tiene mayor poder monopólico que cualquier gobierno estatal. Para la mayoría de la gente es relativamente fácil dejar un estado para reubicarse en otro, pero muy difícil reubicarse en un país diferente. Si un gobierno estatal trata de usar sus facultades fiscales o reguladoras para transferir riqueza de un grupo de ciudadanos a otro, las víctimas podrían optar por mudarse a otro estado. Pero la redistribución tendría que ser enorme antes de que muchas personas trataran de

[1] Por lo que toca al derecho constitucional de la tributación estatal, el tema principal de este capítulo, véase Jerome R. Hellerstein y Walter Hellerstein, *State Taxation*, caps. 8 y 9 (3ª ed., 2001); en cuanto a la jurisdicción de los tribunales federales, también discutida en este capítulo, véase Richard A. Posner, *The Federal Courts: Challenge and Reform,* caps. 2 y 9 (1996).

[2] Por ejemplo, la Corte Suprema ha sostenido que el gobierno federal no puede reclutar a un gobierno estatal dentro del servicio federal. Véase, por ejemplo, National Federation of Independent Businesses *vs.* Sebelius, 132 *S. Ct. Rev.,* 2566, 2601-2609 (2012).

[3] Véanse Michael S. Greve, *The Upside-Down Constitution,* 5, 328 (2012); James A. Gardner y Antoni Abad i Ninet, "Sustainable Decentralization: Power, Extraconstitutional Influence, and Subnational Symmetry in the United States and Spain", 59 *Am. J. Comp. L.,* 491 (2011).

rehacer su vida en otro país, sobre todo porque los Estados Unidos son la nación grande más rica del mundo.

Aunque el monopolio establecido tiende a ser más dañino a nivel federal que estatal, resulta más difícil de establecerse a nivel federal. Tal como lo recalcó Madison al defender la posibilidad de afincar un gobierno republicano de carácter estable en los Estados Unidos, entre mayor y más heterogéneo sea el cuerpo político, mayores serán los costos de transacción derivados de organizar una coalición dominante. Por lo tanto, los costos esperados del monopolio podrían no ser mayores a nivel federal que estatal, una vez que se tome en cuenta la mayor probabilidad del monopolio a nivel estatal. Por supuesto, si los costos esperados son los mismos pero la mayoría de la gente siente aversión por el riesgo, la *desutilidad* esperada del monopolio podría ser mayor a nivel federal. Esto podría explicar el hecho de que la separación de poderes, que vuelve más difícil el establecimiento del monopolio político, sea más elaborada a nivel federal que en la mayoría de los estados.

2. Las deseconomías de escala y la diversidad de enfoques. Tanto en el gobierno como en las empresas privadas, todo intento de dar forma a una organización estrictamente jerárquica ("en forma de U", esto es, unificada o unitaria), en vez de una acoplada de manera más laxa u horizontal ("en forma de M", esto es, de múltiples divisiones), encara desventajas. Con una jerarquía estricta, la información fluye del nivel inferior de la empresa a los estratos más elevados de la administración, y las órdenes fluyen en sentido inverso con base en las decisiones tomadas en la cúpula. Resultará inevitable que la información se pierda o quede tergiversada en su trayectoria de abajo hacia arriba, y, a consecuencia de ello, los gerentes ubicados en el nivel superior fundarán forzosamente sus decisiones en informes que a menudo adolecerán de inexactitudes o de falta de datos; y, de modo semejante, los mandatos tenderán a ser mal comprendidos al circular hacia abajo a través de las sucesivas articulaciones que integran el cuerpo de la unidad en cuestión. La centralización del poder a cargo de las resoluciones también habrá de reducir la competencia, la diversidad y la flexibilidad; las decisiones erróneas serán más costosas debido a que tendrán un carácter vinculante para la empresa considerada en su totalidad; y los errores sobrevendrán con frecuencia porque los administradores situados en la alta esfera no tendrán a su disposición una gama completa de alternativas de entre las cuales escoger, ya que sus subordinados eliminarán, en una operación de filtrado, la mayor parte de éstas en su camino hacia arriba a fin de ahorrarles a sus jefes la experiencia de empantanarse en un océano de datos. En principio, cualquier organización puede evitar problemas de gigantismo y de estructura monolítica adoptando una forma descentralizada, tal como lo llevan a cabo muchas empresas de negocios al integrar diferentes ramas como centros destinados a dar rendimientos por separado. Empero, esto resulta más difícil en el gobierno, pues esos centros no se pueden formar a base de entidades (las

dependencias o agencias gubernamentales) que no venden sus productos en el mercado y, por lo tanto, no generan utilidades. Puesto que ya contamos con gobiernos estatales, puede haber más lógica en la resolución de asignarles funciones en las que el objetivo deseable sea conducir las cosas según un patrón de diversidad o de pequeña escala, en lugar de procurar que se descentralice el gobierno federal.

¿Qué tan lejos debiera seguir avanzando la descentralización en sentido descendente?[4] Deberíamos distinguir entre la descentralización del poder federal y la del estatal. Los estados se encuentran descentralizados; delegan los poderes a entidades subordinadas tales como los condados y las ciudades. El gobierno federal se halla descentralizado de manera similar: pensemos solamente en el sistema judicial federal, que se divide geográficamente en circuitos y, dentro de éstos, en distritos. Lo que sería inviable es, tras abolir los estados, descentralizar al gobierno federal hasta, digamos, el nivel rural. La cantidad de coordinación requerida sería formidable.

3. Las externalidades. Si los beneficios o los costos de una actividad dentro de un estado llegan a quienes no son residentes (la externalidad podría ser el propio gobierno estatal, como veremos), se distorsionarán los incentivos del gobierno estatal. Nuevamente, existen analogías con la organización de un corporativo o empresa. Cada división de una firma tenderá a pasar por alto los efectos que tienen sus formas de actuar sobre las demás divisiones; cada una será renuente a incurrir en costos que beneficien a las demás divisiones pero no lo será a imponer costos sobre las demás. La centralización es una manera de internalizar costos y beneficios a través de una empresa coordinando las divisiones y asegurándose de que estén bien engarzadas. En el caso del gobierno, sin embargo, podríamos pensar que se aplicaría redondamente el Teorema de Coase si sólo dos o tres estados están implicados: si la contaminación del estado A ensucia el aire en el estado B, ¿por qué no negocian A y B la solución que minimice los costos? Los obstáculos son: *1)* el contexto de monopolio bilateral; *2)* la dificultad de ejecutar un fallo legal contra un estado recalcitrante; *3)* la ausencia de incentivos fuertes a cualquier nivel gubernamental para minimizar los costos, y *4)* la dificultad de decidir cómo asignar un pago a un estado entre sus ciudadanos.[5]

Dado que existen tanto costos como beneficios derivados de la centralización, una que sea eficiente generalmente tendrá a la vez elementos propios del control central y una autonomía a nivel de las divisiones; en otras palabras, habrá en ella cierta jerarquía, pero no en forma excesiva, y sus divisiones

[4] Véanse Maria Teresa Balaguer-Coll, Diego Prior y Emili Tortosa-Ausina, "Decentralization and Efficiency of Local Government", 45 *Annals of Regional Sci.,* 571 (2010); Oded Hochman, David Pines y Jacques François Thisse, "On the Optimal Structure of Local Governments", 85 *Am. Econ. Rev*., 1224 (1995).

[5] Véase Barry R. Weingast, "The Economic Role of Political Institutions: Market Preserving Federalism and Economic Development", 11 *J. Law, Econ. & Organization,* 1 (1995).

presentarán un perfil semiautónomo en vez de autónomo.[6] Y esto es lo que observamos en nuestro sistema federal. Los estados gozan de considerable autonomía en lo concerniente a la tributación y la regulación (incluida la acreditación profesional); a la administración de las escuelas y de las prisiones; a las carreteras y otra clase de infraestructuras; a la aplicación rigurosa del derecho penal; y al matrimonio, el divorcio, la custodia y otros asuntos relativos a la familia; no obstante, el Congreso, la Corte Suprema y el presidente disponen de un considerable poder para neutralizar las regulaciones. La Constitución empodera al Congreso para regir en materia de comercio interestatal y extranjero; y, según veremos, la Corte Suprema, a través de la interpretación de la cláusula que confiere ese poder al Congreso, ha prohibido a los estados imponer aranceles y otros impedimentos al comercio y los viajes interestatales incluso cuando al Congreso no se le posibilita emprender acciones.

En virtud de que los estados disponen de cierto grado de autonomía, funcionan, de modo muy semejante a las divisiones de una compañía de software o una farmacéutica, como laboratorios para la experimentación (social). Una serie de políticas aplicadas en un estado, de resultar exitosas, pueden ser copiadas por otros. Asimismo, las personas pueden transitar por estados distintos de acuerdo con sus preferencias. El derecho de trasladarse a otro estado sirve de complemento al poder de voto en el esfuerzo de controlar la forma de proceder de los funcionarios gubernamentales.

XXVII.2. La respuesta a las emergencias: Katrina (o Sandy) y el terrorismo

Los funcionarios públicos que por su posición están más compenetrados de un problema cuentan con la información más precisa y también pueden actuar con mayor prontitud al respecto. No nos daríamos por bien servidos si contáramos (únicamente) con un Departamento Federal de Bomberos, de modo que en el caso de un incendio local de nivel 2 el jefe del departamento local tendría que informar a Washington y obtener el permiso para combatir el fuego. Éste es el punto que esgrimen quienes creen que la Agencia Federal para el Manejo de Emergencias (FEMA), aun cuando la presida una jefatura competente en su ejercicio del liderazgo, no debería estar encabezando una respuesta de emergencia a las catástrofes. Con todo, no todas las catástrofes son de carácter local. Ni tampoco, dada la movilidad propia de los que responden a ellas, es razonable que cada localidad invierta en los recursos necesarios para lograr autosuficiencia en el manejo de una emergencia, independientemente de su magnitud. Supongamos, como en el caso del huracán Katrina, o en el del derrame masivo de petróleo ocasionado por la explosión de una

[6] Véase Steven G. Calabresi y Nicholas Terrell, "The Number of States and the Economics of American Federalism", 63 *Fla. L. Rev.,* 2, 43-44 (2011).

plataforma petrolera en el golfo de México el mes de abril de 2010, que la catástrofe hace presa simultáneamente de un gran número de ciudades y pueblos en varios estados. En la medida en que resulta óptima una respuesta bien coordinada, y dado que los costos de transacción son particularmente altos en una situación de emergencia, no tiene ningún sentido dejar la respuesta en manos de los gobiernos estatales y locales. Cada estado buscará llevar al máximo su capacidad de hacer frente al daño que le ha causado el desastre, y cada localidad procederá de la misma manera ante el daño que se le ha causado, reduciendo los costos y beneficios que tengan sus acciones para otros estados y localidades. Más aún, a semejanza de lo acontecido en una situación militar en la que no sabe uno por dónde atacará el enemigo, una respuesta efectiva a una emergencia exige que se mantengan reservas listas para desplegarse en el lugar donde se cierne la amenaza, y es sin duda el centro el que debe ejercer control sobre dichas reservas.

Asimismo, si el gobierno federal le indica a una ciudad que no recibirá asistencia alguna en el caso de sobrevenir una catástrofe, por grande que sea ésta, la ciudad invertirá con creces en su capacidad de encarar un desastre. Supongamos que una nación contara sólo con dos ciudades, que el costo de responder a un desastre promedio fuese de 1 y a un cataclismo de 20, pero que la posibilidad de que ocurriesen dos cataclismos al mismo tiempo fuese cercana a 0. Entonces, si el gobierno federal se rehusara a prestar auxilio en lo tocante a desastres locales, no importa cuán destructivos, las dos ciudades podrían incurrir en un costo total de prevención de 40, en tanto que, si el gobierno federal invirtiera en los recursos necesarios para la capacidad de respaldo suficiente, el costo total de prevención descendería a 22 (1 + 1 + 20).

Haciendo a un lado los desastres naturales, dirijamos ahora la vista hacia los que provoca la intervención humana. El Departamento de Seguridad Nacional de los Estados Unidos (DHL) les proporciona a las ciudades más de 500 millones de dólares para la adopción de medidas de antiterrorismo, principalmente bajo la forma de capacidad física para responder a ataques terroristas en lugar de sufragar gastos recurrentes tales como los de salarios del personal involucrado en la respuesta en cuestión. Abandonadas a sus propios medios, las ciudades podrían optar —más allá del criterio de eficiencia— por gastarse el dinero en salarios y no en bienes de capital producidos en otra parte, dado que los empleados de la alcaldía son habitualmente residentes locales.

¿Debiera corresponder al gobierno federal, y no solamente a los estados y municipios, gastar dinero con la finalidad de proteger a la ciudad de ataques terroristas? Indudablemente, sí. Los ataques del 11 de septiembre de 2001 demostraron que toda acometida sobre una ciudad o sobre cualquier otro blanco de envergadura ha de tener por fuerza repercusiones a escala nacional. Empero, ¿debiera el gobierno ocuparse de financiar medidas defensivas tomadas por las ciudades, o más bien estaría obligado a gastar directamente

su dinero en tales medidas? Probablemente le corresponde hacer ambas cosas. Los representantes locales saben mejor cuáles son los puntos débiles de sus ciudades y conocen mejor la ruta para fortalecerlos. No obstante, ocurre a menudo, como sucedió en el caso de los bombardeos del Maratón de Boston, que a duras penas saben algo sobre los planes, métodos, blancos preferidos, etc., de los terroristas. Y pueden utilizar el dinero federal sencillamente para remplazar los gastos que de otra manera habrían tenido que hacer en lo concerniente a las medidas antiterroristas. Si el remplazo es completo, entonces el programa federal —junto con todo lo que implica en materia de papeleo, revisión por pares, la necesidad de lidiar con las controversias políticas y la de aplacar las preocupaciones en el sentido de que la concesión de subvenciones federales a distintas localidades con base en la "calidad" de las cartas de presentación de proyectos está meramente recompensando el esfuerzo de epistológrafos habilidosos— no vendría a constituir una medida de seguridad; sería sólo un subsidio federal de carácter general brindado a un gobierno local. Con todo, este problema podría ser superado al menos parcialmente recurriendo a donaciones por contrapartida.

XXVII.3. Los tribunales federales y la aplicación de la ley federal

Aunque la jurisdicción de diversidad de los tribunales federales se justifica tradicionalmente por referencia a la presunta hostilidad hacia quienes no son residentes, hay una explicación económica, por lo menos para una parte de la jurisdicción, que no se relaciona con la xenofobia. Supongamos que un residente del estado A tiene una colisión con un residente del estado B en B, y que el residente de B demanda al residente de A en un tribunal estatal en B. Si el demandante gana, los beneficios serán para B; si el demandado gana, los beneficios serán para A; por lo tanto, el tribunal podría inclinarse hacia el demandante.

Tal inclinación es menos probable si las partes tienen una relación preexistente. Supongamos que el juicio surge de un contrato entre un residente de A y un residente de B. Si se sabe que los tribunales de B favorecen a los residentes, el contrato requerirá que cualesquiera disputas surgidas del contrato no se resuelvan en un juicio en B, o que el residente de B tendrá que compensar al residente de A en el precio del contrato o alguna otra cláusula por soportar el riesgo adicional de la litigación en B. Los tribunales de B no ganarán nada a largo plazo por inclinarse en favor de los residentes.

La Ley Federal de Reclamaciones de Cuasidelitos permite que quienes resulten perjudicados por cuasidelitos cometidos por empleados federales demanden a los Estados Unidos, pero sólo en tribunales federales. La justificación económica del confinamiento de estos juicios a los tribunales federales es similar a la de la jurisdicción de diversidad. Si la camioneta de correos

choca con un residente del estado A, quien podrá demandar al Servicio Postal en un tribunal estatal de A, ese tribunal podría inclinarse en su favor, sabiendo que si el demandante gana se dispersarán los costos del Servicio Postal por todos los Estados Unidos en lugar de concentrarse en A.

Gran parte de la legislación penal federal (que se aplica exclusivamente en tribunales federales) puede explicarse como una respuesta al problema de las externalidades interestatales. Supongamos que una pandilla de delincuentes opera en muchos estados. Cada estado tendrá cierto incentivo para someterla a la justicia, pero menos que el que tendría si los costos totales de las actividades de la pandilla se sufragaran dentro del estado. Por supuesto, si cada estado realiza algunas investigaciones, los recursos agregados gastados en la investigación podrían ser tan grandes como si una sola entidad fuese responsable de la investigación. Pero los recursos no se emplearán tan eficientemente porque será costoso coordinar las actividades de los diferentes investigadores de los distintos estados.

Una clase de externalidad diferente justifica también un papel federal en la aplicación de la ley penal. Si una persona defrauda a un banco cuyos depósitos están asegurados por la Corporación Federal de Seguro de Depósitos, una parte de los costos del fraude podría ser soportada por los accionistas y otras personas del estado, pero la mayor parte de los costos será sufragada por los contribuyentes federales, dependiendo, sin embargo, de cómo se determinen las tasas de seguros de depósitos. Si las tasas están "clasificadas según la experiencia", es decir, si se les permite variar de acuerdo con el grado de riesgo del asegurado en particular, no existe ninguna externalidad. No obstante, de hecho no existe ninguna clasificación según la experiencia del seguro de depósito federal.

En cambio, los costos de un robo de banco se harán sentir principalmente en el estado donde ocurra el robo, porque los costos no pecuniarios en términos del temor (y a veces en términos de las lesiones) dominan sobre los costos financieros de ordinario moderados de un robo de banco, y aquellos costos se sufren localmente. Por lo tanto, no debiera sorprendernos que, mientras que el robo de banco es, como el fraude bancario, un delito federal tanto como estatal, la mayoría de los robos bancarios se procesen de acuerdo con la ley estatal antes que la federal: lo contrario del patrón observado en los casos de fraudes bancarios.

Cierta jurisdicción penal federal puede explicarse por referencia a la observación hecha en la sección anterior de este capítulo en el sentido de que los monopolios del poder político se establecen más fácilmente a nivel estatal que a nivel federal. Los procesamientos penales federales de funcionarios públicos locales corruptos aprovechan la relativa incorruptibilidad de los funcionarios federales —derivada de los mayores costos de la corrupción de una dependencia federal— a fin de disminuir la corrupción a nivel local. De la relativa incorruptibilidad que prevalece entre los funcionarios federales da cuenta la

relativa escala y poder de la aplicación de la ley federal *versus* la estatal (y la de otras agencias: el punto es general). Entre más grande sea el alcance y poder de una dependencia y, por consiguiente, mayores sean los costos sociales en el caso de que sea corrupta, más grandes serán los esfuerzos para prevenir que se corrompa (se realizará una supervisión más estricta, etc.). Asimismo, los servidores civiles federales reciben generalmente una retribución mayor que la de sus contrapartes estatales, y, por lo tanto, tienen más que perder si se descarrían y los atrapan. Igualmente, debido a que las agencias federales se distinguen por tener una jurisdicción geográfica más amplia, resulta más difícil para sus empleados establecer el tipo de vínculos locales que generan una base donde la confianza y, concomitantemente, la colusión pueden prosperar.

En décadas recientes, a medida que ha crecido la preocupación pública por el delito, se han promulgado más y más leyes penales federales que duplican las prohibiciones penales estatales sin servir a un interés federal claro. (En la última sección de este capítulo examinaremos un ejemplo particularmente destacado.) La razón principal es que el derecho penal federal es más favorable para el procesamiento que el derecho penal estatal. Los "federales" emplean en mayor medida la detención anterior al juicio; hay una autoridad más amplia para la utilización de grabaciones; los jueces federales tienen la mano más dura en los procesos penales; los grandes jurados, que han evolucionado de ser una protección para los acusados hasta convertirse en potentes instrumentos de investigación para los fiscales, se utilizan más en el sistema federal; la doctrina de la conspiración es más expansiva en el derecho penal federal; y las sentencias son más largas y se ha abolido la libertad bajo palabra (si bien se ha añadido al final de la mayoría de las sentencias penales federales un sustituto denominado "liberación supervisada"). La Corte Suprema, al imponer a los estados casi todas las salvaguardas procedimentales otorgadas por la Carta de Derechos de los Estados Unidos a los acusados de delitos federales en nombre del debido proceso, ha colaborado para debilitar (sin querer) al derecho penal estatal en relación con el federal.

La razón más importante para explicar la diferencia de poder existente entre el cumplimiento del derecho estatal y el federal, con todo, es que el derecho penal federal tiene que ser más efectivo a fin de proteger a la gente contra las actividades delictivas suficientemente graves para tener un impacto nacional potencial, como ocurre con el sabotaje, el terrorismo, la falsificación y la insurrección. Este poder queda entonces disponible para afrontar los delitos locales cuando surge la necesidad de una aplicación más eficaz de la ley. Una explicación final de que el derecho penal federal sea más efectivo que el estatal consiste en que los recursos del gobierno federal, medidos en relación con la cantidad de delitos cometidos a lo largo de la federación, arrojan una suma mucho mayor a la que constituye la razón de los recursos estatales a la cantidad de delitos estatales. El gobierno federal, a diferencia de los gobiernos es-

tatales, no se halla sometido a presión para sacar a los delincuentes convictos de la prisión tras haber sido encarcelados por un periodo breve a fin de economizar en materia de gastos penitenciarios.

Un buen ejemplo de las áreas del derecho civil federal que pueden justificarse por referencia a las externalidades interestatales es el de la quiebra. Si la quiebra fuese un campo del derecho estatal, los deudores podrían eludir sus deudas mudándose a los estados que tuvieran leyes de quiebras favorables al deudor, y así se debilitaría el mercado nacional del crédito; sería difícil que los acreedores impidieran esto mediante una redacción cuidadosa del instrumento crediticio. Sin embargo, lo que resulta extraño desde esta perspectiva, el Código de Quiebras, al permitir que el deudor opte por las exenciones impuestas al embargo en su estado, conserva un incentivo para que los deudores traten de eludir sus deudas mudándose a estados que tengan, por ejemplo, una generosa exención para la casa familiar.

XXVII.4. Tributación estatal: impuestos específicos

Todos los estados tienen un incentivo para crear impuestos cuyo peso recaiga en la mayor medida posible sobre los residentes de otros estados. Tal tributación desvía al Estado de la búsqueda de métodos de tributación que maximicen los valores de eficiencia y distribución para el conjunto de la nación, y también conduce a un aumento de los gastos gubernamentales al permitir que el estado externalice los costos de sus servicios públicos. Los estados podrían emplear también la tributación no para incrementar su recaudación, sino para proteger a los productores del estado o a otros vendedores contra la competencia de quienes no sean residentes. Tal tributación perjudicará a los consumidores residentes al igual que a los vendedores no residentes de los bienes de fuera del estado que están gravados, y, en efecto, la pérdida de los consumidores residentes superará de ordinario la ganancia de los vendedores residentes.[7] Pero, dado que los intereses especiales obtienen con frecuencia una legislación que reduce la eficiencia, no debiera sorprendernos el hecho de que, en los estados como en las naciones, a veces se establezcan impuestos que discriminan contra los importadores. Tal impuesto distorsiona la distribución geográfica óptima de la empresa. La producción que podría realizarse más eficientemente en el estado A se realizará en B si el impuesto de B a las importaciones provenientes de A supera la ventaja de costos de los productores de A.

[7] El impuesto aumenta el precio del bien importado en relación con el del bien local y así induce a los consumidores a preferir el último. La sustitución beneficia al vendedor local a la vez que perjudica al consumidor. Pero no todos los consumidores hacen la sustitución. Algunos continúan comprando el bien importado al precio más alto, y su pérdida no es captada por el vendedor residente. El análisis es similar al del monopolio: los consumidores pierden más de lo que gana el monopolista (véase antes IX.3).

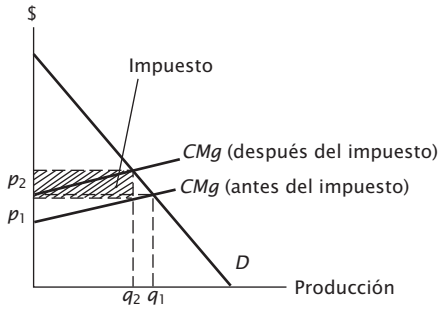
Debemos vigilar entonces dos cosas separadas: un impuesto estatal que exporta los costos del gobierno estatal, y un impuesto estatal que protege a los productores locales contra la competencia de los productores de fuera del estado. Sin embargo, debe tenerse presente que: *1)* un impuesto que discrimine en favor de las empresas de fuera del estado causará la misma distorsión de las ventajas comparativas geográficas que un impuesto que discrimine en su contra, y *2)* no hay, empero, ninguna discriminación, en contra de los consumidores no residentes o de los productores no residentes, si la diferencia de las cargas tributarias refleja una diferencia en el costo de los servicios provistos por el estado a los diferentes grupos de contribuyentes.

La Corte Suprema tiene una autoridad formal adecuada en los términos de la cláusula de comercio y la cláusula de los privilegios e inmunidades de la Constitución original, y las cláusulas del debido proceso y la protección igual de la Decimocuarta Enmienda, para impedir que los estados creen impuestos que tengan una de las características cuestionables o ambas que acabamos de identificar (gravar a los consumidores no residentes y excluir a los productores no residentes). Ha ejercido esta autoridad para prohibir los impuestos claramente extraterritoriales, así como los aranceles fijados a los bienes importados de otros estados, pero a menudo no ha impedido que los estados restablezcan tales impuestos y aranceles bajo nombres que ocultan el verdadero efecto económico.

Si Montana cobrara un impuesto a la venta del carbón en Illinois, sería clara la invalidez de tal gravación; pero un impuesto a la extracción de carbón de las minas de Montana (un impuesto de extracción) es claramente válido. El tratamiento doble no tiene sentido económico. Si la demanda del carbón de Montana es relativamente inelástica y los consumidores finales de este carbón son principalmente no residentes, el impuesto, suponiendo que es proporcional a la cantidad producida, será pagado principalmente por no residentes bajo la forma de precios mayores del carbón, como se muestra en la gráfica XXVII.1.

Podría sostenerse que el impuesto a la extracción es en efecto una regalía a la que debiera tener derecho el estado como dueño original de los recursos naturales que se encuentren dentro de sus fronteras.[8] Como veremos cuando discutamos las transferencias de agua entre cuencas, es cierto que se promueve la eficiencia otorgando al estado la propiedad de los recursos naturales en los que los individuos privados o las empresas no hayan obtenido todavía derechos de propiedad; ese otorgamiento inicial es el primer paso para la creación de mercados de los recursos. Pero si el estado conserva la propiedad de todos los derechos sobre un recurso o los vende todos a un solo comprador, el resultado —si no hay buenos sustitutos, como se supone en nuestro ejemplo del carbón— será un monopolio. El monopolio no es inherente a la

[8] Debe señalarse un problema fáctico de este argumento: excepto en Texas y Alaska, el propietario original y actual de la mayoría de los recursos naturales es el gobierno federal, antes que los estados.



GRÁFICA XXVII.1

propiedad originalmente estatal de "sus" recursos naturales; el estado puede dispersar los derechos en una forma suficientemente amplia para asegurar la competencia; un requerimiento de que se efectúe realmente el desvío en el uso a fin de establecer la propiedad de los derechos sobre el agua limita la cantidad de ésta que cualquier usuario pueda reclamar y, así, impide que se delineen reclamaciones desproporcionadas que pudiesen conferir un poder monopólico. Pero un impuesto a la extracción, cobrado sobre un recurso que carece de buenos sustitutos producidos en otros estados, es monopólico. El impuesto eleva el precio por encima del nivel competitivo y disminuye la producción por debajo de tal nivel, y genera recaudaciones por encima del nivel competitivo, aunque las reciba el estado antes que los tenedores de derechos.

O supongamos que un estado impone gravaciones onerosas a todos los camiones que utilizan sus carreteras, teniendo conciencia de que casi todo el tráfico de camiones encuentra su origen y su término en otros estados. Estaría el estado sacando provecho del monopolio de una locación para derivar los costos de servicios públicos no relacionados con el mantenimiento de la red de carreteras hacia los proveedores y consumidores ubicados en otros estados.

Se trata de un problema muy añejo, el cual se ilustra, por ejemplo, en la obra clásica de la literatura alemana *Michael Kohlhaas*, novela corta de Heinrich von Kleist. Cuando se publicó el libro, en 1810, lo que ahora es Alemania se dividía entre cientos de estados independientes. Un camino, digamos, de Múnich a Berlín atravesaría muchas fronteras, y cada estado en el que desembocara el camino podría cobrar derechos de peaje como una condición para autorizar el ingreso. La tributación estaría limitada no por el costo que impusiera al estado el viajero, en términos de desgaste de los caminos y de congestionamiento del tráfico, sino por el costo que representara para el viajero utilizar una ruta alterna menos directa. A semejanza de lo que sucedía en la Alemania de comienzos del siglo XIX, en los Estados Unidos la red eléctrica solía estar dividida entre cientos de servicios independientes, cada uno de los cuales cobraba por separado una tributación por el derecho a enviar electricidad a través de la porción de la red que le correspondía. Los múltiples cobros

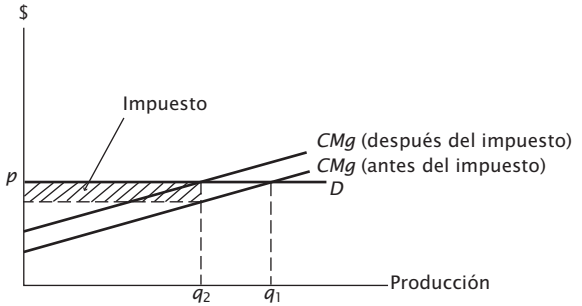
que se fijaban a lo largo de la transmisión a larga distancia desalentaban la puesta en marcha de ésta. La Comisión Federal Reguladora de la Energía (FERC) promovió la creación de las Organizaciones Regionales de Transmisión (RTO) como una fórmula para eliminar esos monopolios de localidad: exigía que las RTO abarcaran de modo coherente regiones geográficas enteras y que cada RTO cobrara una sola cuota por el uso de la red en su integridad.[9]

No todos los impuestos específicos generan los efectos de mala asignación asociados al monopolio. Si hubiera buenos sustitutos para el carbón de Montana, de modo que la curva de demanda de la gráfica XXVII.1 fuese horizontal, el impuesto específico sólo disminuiría la producción de ese carbón y el valor de los campos carboníferos de Montana, conduciendo a un aumento compensatorio de la cantidad de producción y los valores de la tierra en otra parte. Sería un impuesto a la renta económica (véase la gráfica XXVII.2). Sería una forma de la tributación responsable y, así, una gravación responsable y eficiente, siempre que los propietarios de los campos carboníferos de Montana fuesen residentes de ese estado, porque entonces los residentes del estado que crea el impuesto serían los contribuyentes reales y no meramente nominales.

Ocurre un caso intermedio cuando, aunque sea inelástica la demanda del recurso gravado, porque no hay buenos sustitutos en otra parte, hay un mercado sustancial dentro del estado, de modo que la carga mayor del impuesto recae sobre los residentes. Todavía podría objetarse que el impuesto sería monopólico en sus efectos, pero la objeción no sería vigorosa. Virtualmente todos los impuestos tienen efectos de mala asignación porque las recaudaciones demandadas por el gobierno moderno no pueden empezar a obtenerse mediante impuestos a las rentas económicas exclusivamente. Mientras la incidencia del impuesto sea local, no hay ninguna o n especial; porque entonces existen los frenos políticos habituales sobre el nivel del impuesto.

La Corte Suprema ha sido más crítica de los impuestos estatales a la importación, es decir, de la exclusión de los productores no residentes por oposición al gravamen de los consumidores no residentes. Algunas decisiones anteriores sostenían que un estado no podría cobrar su impuesto general a las ventas sobre las ventas hechas a sus residentes por vendedores de fuera del estado. Si se pregunta cómo podría discriminar la recaudación uniforme de un impuesto a las ventas contra tales vendedores, considere dos estados, uno que recauda primordialmente mediante un impuesto a las ventas (estado A), el otro mediante un impuesto a la propiedad (B). Dado que el valor de la propiedad comercial es principalmente el valor presente de sus ganancias netas anticipadas, el ingreso neto derivado de las ventas hechas a los residentes del estado A por una empresa ubicada en el estado B se capitalizará en el impuesto a la propiedad pagado por la empresa a B, de modo que la empresa

[9] Véase Illinois Commerce Commission *vs*. Federal Energy Regulatory Commission, 2013 WL, 2451766 (7 de junio, 2013).



GRÁFICA XXVII.2

está pagando a B, en efecto, un impuesto sobre sus ventas en A. Si la empresa está sujeta también al impuesto a las ventas del estado A, pagará más impuestos que sus competidores en A (quienes pagan sólo el impuesto a las ventas), mientras que no recibe mayores servicios gubernamentales. Este diferencial impositivo, que no se relaciona con ninguna diferencia de los costos, discrimina en contra del vendedor de fuera del estado, aunque el efecto podría disminuir por el hecho de que el impuesto a las ventas del estado A disminuye el valor de la propiedad de la empresa y, por ende, la cantidad de impuestos que paga a B.[10]

No obstante, la Corte ha sostenido un instrumento claramente evasivo conocido como un impuesto al uso compensatorio. Éste es un impuesto, igual al impuesto a las ventas, sobre los bienes que compran los residentes del estado que grava pero sobre los cuales no pagan el impuesto a las ventas. En nuestro ejemplo, los residentes del estado A que compraron a la empresa en B tendrían que pagar a A un impuesto igual al impuesto sobre las ventas que habrían pagado si hubiesen comprado los mismos bienes a un vendedor en A. La Corte ha permitido también que los estados obliguen al vendedor de fuera del estado a recaudar el impuesto al uso compensatorio, lo que completa la identidad funcional de los dos impuestos.

[10] Disminuido, pero no eliminado. Supongamos que el impuesto a las ventas del estado A es de 3 centavos por dólar, la ganancia normal del vendedor de fuera del estado es de 10 centavos por dólar y el impuesto predial del estado B equivale a 30% de la ganancia normal del vendedor, de modo que también es de 3 centavos. Si no se cobra el impuesto a las ventas del estado A, el vendedor de fuera del estado obtiene una ganancia neta de 7 centavos por dólar. Si se cobra el impuesto, la ganancia del vendedor antes del cobro del impuesto predial es de sólo 7 centavos. El impuesto predial es 30% de esta cantidad, o sea 2.1 centavos de dólar, de modo que la carga impositiva total del vendedor es de 5.1 centavos, dejando una ganancia de sólo 4.9 centavos. Por lo tanto, si los vendedores del estado A sólo están sujetos al impuesto a las ventas, el cobro de ese impuesto al vendedor de fuera del estado les otorga una ventaja competitiva no relacionada con la eficiencia, pues estamos suponiendo que el vendedor de fuera del estado no recibe servicios gubernamentales mayores aunque sus impuestos sean más de 40% mayores a los de los vendedores locales. Pero está claro también que la carga fiscal adicional del vendedor de fuera del estado (2.1 centavos de dólar) es menor a la suma de los dos impuestos que lo gravan.

Es posible que la razón de que la Corte no haya "visto más allá" de los impuestos de extracción y de uso compensatorio —el primero que grava a los consumidores no residentes, el segundo que discrimina a los productores no residentes— es que hasta hace pocos años planteaba la cuestión en tales casos en términos de si el estado había establecido un impuesto "en" el comercio interestatal. El petróleo tal como sale del suelo y el bien que ha sido recibido por el comprador residente están seguramente presentes, en un sentido físico, dentro de las fronteras del estado. Pero es irrelevante desde un punto de vista económico que un impuesto se grave "en" el comercio interestatal o donde el sujeto nominal del impuesto se ubique físicamente. Los problemas económicos son a cuánto asciende el monto del impuesto soportado en última instancia por los no residentes y si el efecto del impuesto es el de aumentar, sin una justificación basada en los costos mayores de los servicios gubernamentales, los precios de los bienes importados por comparación con los de los bienes del estado.

En algunos casos, el hincapié en si un impuesto se cobra "sobre" el comercio interestatal ha otorgado rebajas tributarias injustificadas para los vendedores de fuera del estado, creando así un incentivo opuesto pero igualmente poco apropiado para sustituir los bienes y servicios intraestatales por los interestatales. Supongamos que una empresa ubicada en un estado que recurre primordialmente a un impuesto a las ventas para aumentar sus recaudaciones realiza la mayor parte de sus ventas a residentes de otros estados que recurren primordialmente al impuesto a la propiedad para aumentar sus recaudaciones. Si se prohíbe que el estado de residencia de la empresa aplique su impuesto a las ventas interestatales de la empresa, porque tales ventas corresponden al comercio interestatal, la empresa pagará impuestos menores a los de los competidores en los estados donde vende que hagan sobre todo ventas *intraestatales* (ya que la empresa proveniente de otro estado no paga impuestos a la propiedad, mientras que las asentadas en el estado sí los pagan), aunque no reciba menos servicios gubernamentales que tales competidores.

Los dos criterios económicos de los impuestos estatales sospechosos (a saber, si el impuesto recae primordialmente sobre los no residentes y si tiene un efecto distorsionador sobre lo que ya constituía comparativamente una ventaja geográfica) podrían entrar en conflicto en casos particulares. En el ejemplo de un impuesto a las ventas sobre una empresa que vende principalmente en otros estados, el primer criterio implicaría una política de prohibición de la tributación de las ventas interestatales de la empresa por el estado del domicilio, porque la carga del impuesto recaería primordialmente sobre los no residentes. Pero el segundo criterio nos llevaría a aprobar tal impuesto, de modo que los impuestos pagados por una empresa interestatal no serán menores a los de los pagados por una empresa local debido a una diferencia de sus posiciones antes que de la cantidad de servicios gubernamentales que reciben.

¿Podrán conciliarse los criterios? Un enfoque podría consistir en: *1)* permitir que todos los estados establezcan un impuesto general a las ventas aplicado indiscriminadamente a las ventas interestatales e intraestatales realizadas por los vendedores del estado, pero: *2)* prohibir los impuestos a productos particulares que sean primordialmente bienes de exportación cuando los productos de importancia equivalente vendidos primordialmente en el mercado local no estén gravados en la misma medida. De acuerdo con este enfoque, los impuestos a la extracción serían dudosos,[11] y se eliminaría el último vestigio de justificación de los impuestos al uso compensatorio porque el vendedor en el exterior del estado en ese caso estaría sujeto al impuesto a las ventas en su estado de residencia aun aplicado a sus ventas en el exterior del estado.

XXVII.5. Tributación estatal: impuestos prediales y al ingreso corporativo

La Corte Suprema ha tendido a tratar los impuestos prediales como inherentemente locales. Lo serían si fuesen sólo impuestos a la renta de la tierra y todas las tierras fuesen de propiedad local. Pero, en virtud de que la renta de la tierra es sólo un componente (y uno limitado) de la base del impuesto predial (véase antes XVIII.4), un estado que grave la tierra de un ferrocarril, por ejemplo, probablemente la gravará en proporción a los ingresos del ferrocarril en lugar de tratar de limitar el impuesto a la renta de la tierra, el cual por añadidura puede resultar difícil de calcular. El ferrocarril tratará el impuesto predial basado en los ingresos como uno específico y lo trasladará en la mayor medida posible a los clientes del ferrocarril, la mayoría de los cuales serán no residentes, elevando sus tarifas. (Que pueda el ferrocarril llevar a cabo esta acción, y en qué medida lo logre, dependerán de las elasticidades de la demanda y la oferta. Véanse las gráficas XVIII.1 y XVIII.2 en el capítulo XVIII.) En el grado en que el impuesto grave las rentas, la carga será soportada de nuevo primordialmente por no residentes: los accionistas del ferrocarril.[12]

Un peligro equivalente acecha a los impuestos estatales al ingreso corporativo. La carga de tal impuesto se reparte normalmente entre los consumidores y los accionistas (véase antes XVIII.5). En el caso de una corporación que realiza negocios en una variedad de estados, la mayoría de sus clientes y accionistas no serán residentes del estado que grava. Por lo tanto, es grande la tentación de los estados por cobrar fuertes impuestos al ingreso corporativo de las corporaciones multiestatales. Podría pensarse que no hay peligro de discriminación mientras cada estado tenga una tasa impositiva uniforme para todas las

[11] Si es cierto que son en su mayor parte impuestos a la exportación.

[12] Esto supone que la tierra era propiedad del ferrocarril antes de que se previera el impuesto. Si el ferrocarril compró la tierra después de que se aplicara el impuesto, la carga tributaria se capitalizaría en el precio de venta y, de este modo, sería soportada por el vendedor; el intento de exprimir al ferrocarril no tendría éxito.

corporaciones que hacen negocios en él, tanto las locales como las multiestatales. El problema, un problema clásico del costo conjunto (véase antes XIII.5), es que de ordinario resulta imposible asignar una parte particular del ingreso de una empresa multiestatal a un estado particular. Si una empresa tiene su planta manufacturera en un estado, sus vendedores en otro y sus oficinas corporativas en un tercero, sus costos y sus recaudaciones y, por ende, su ingreso, son el producto conjunto de sus actividades en cada uno de los tres estados. Dado que no hay un medio racional para repartir el ingreso de tal empresa entre los diferentes estados en los que opera, no resulta sorprendente que la Corte Suprema haya permitido que los estados escojan, dentro de límites amplios, cualquier fórmula de reparto que atribuya la mayor parte posible al estado que grava. Pero si un estado donde un ferrocarril tenga muchas vías pero corra relativamente pocos carros reparte el ingreso por el kilometraje de las vías, mientras que un estado donde el ferrocarril tiene relativamente pocas vías pero corre muchos carros reparte el ingreso de acuerdo con el número de carros, el impuesto al ingreso combinado será mayor que el de cualquier par de empresas similares pero locales, una de las cuales se ubica en el primer estado y la otra en el segundo. El esfuerzo de ambos estados por exportar sus cargas impositivas hará que la empresa multiestatal se vea obligada a pagar un impuesto mayor que si operara sólo en un estado, aunque la empresa no requiera mayores servicios gubernamentales. El resultado es la creación de un incentivo ineficiente para hacer negocios en el menor número de estados posible.

Supongamos que, a fin de determinar el ingreso gravable de una corporación en un estado, el estado multiplica el ingreso total de la corporación por el promedio de las tres razones siguientes: la razón del gasto en nóminas en el estado de la corporación a su gasto total en nóminas, la razón del valor de su propiedad en el estado al valor del total de sus propiedades, y la razón de sus ventas en el estado al valor de sus ventas totales. El impuesto que el estado fija a este ingreso calculado está compuesto en realidad por tres impuestos, cada uno de los cuales tiene su propia incidencia: la de un impuesto a la nómina, la de un impuesto a la propiedad y la de un impuesto a las ventas.

Supongamos que una corporación tiene ingreso de inversión. Podría ser una empresa manufacturera pero con reservas en efectivo sobre las que gana intereses. ¿Debiera repartirse este ingreso, para efectos del impuesto estatal al ingreso corporativo, entre los estados donde opera la empresa, y, en su caso, cómo? La regla es que no se reparta sino que se asigne en su totalidad al domicilio comercial de la corporación.[13] Debiera ser repartido, tal vez en proporción al ingreso de manufacturación ganado en cada jurisdicción tributaria. Si es asignado al domicilio comercial, esto puede darle a la corpora-

[13] Allied-Signal, Inc. *vs.* Director, Division of Taxation, 504 U.S., 768 (1992).

ción un fuerte incentivo para establecer el domicilio en la jurisdicción que tenga el nivel más bajo de tributación, independientemente de la eficiencia.[14]

XXVII.6. Otras limitaciones a la regulación estatal del comercio interestatal

El impuesto a la extracción es un ejemplo de la regulación local no discriminatoria que, sin embargo, es objetable debido a que desplaza los costos gubernamentales del estado donde se incurrió en ellos a otros estados, cuyos ciudadanos no tienen voz ni voto en lo que respecta a la decisión del estado de imponer la tributación. Aun cuando la Corte Suprema se ha mostrado renuente a reconocer esta especie de exportación tributaria como una carga sobre el comercio interestatal, ha invalidado otra clase de regulaciones locales no discriminatorias sin siquiera hacer un intento por establecer un equilibrio entre el beneficio local y la carga de fuera del estado (o extranjera): cuando un estado hace la tentativa de regular actividades en otro estado. En Healy *vs.* Beer Institute,[15] Connecticut había aprobado una ley de "afirmación del precio" que exigía a los fabricantes cerveceros comprometerse a que los precios que cobrarían por la cerveza en el estado no fueran más elevados al momento de hacer el envío a su destino que los precios más bajos que cobraban en cualquier estado vecino. No había discriminación alguna a favor de los cerveceros de Connecticut, puesto que en realidad no había ahí fabricantes de cerveza. Aun así, la Corte Suprema invalidó la ley. Un cervecero podía vender cerveza en Nueva York y Connecticut y cobrar un precio más alto en el segundo porque las personas de este estado eran aficionadas a su cerveza en mayor medida que los neoyorquinos. Al confrontar la ley de afirmación del precio de Connecticut y considerar a esta entidad como su mercado más valioso, el cervecero podría decidir elevar su precio en Nueva York hasta el nivel de su precio en Connecticut en vez de reducir su precio en este último. Así pues, se estaría dando la situación de un estado que regula los precios de otro estado, comoquiera que fuese indirectamente. El comercio se vería obstaculizado si los estados pudieran regular las actividades comerciales en otros estados. La Corte estableció que la ley de Connecticut violaba la cláusula de comercio.[16]

[14] Véase Walter Hellerstein, "State Taxation of Corporate Income from Intangibles: Allied-Signal and Beyond", 48 *Tax L. Rev.*, 739 (1993).

[15] 491 U.S. 324 (1989).

[16] Véase también Midwest Title Loans, Inc. *vs.* Mills, 593 F.3d, 660 (7th Cir., 2010), que da el siguiente ejemplo hipotético: "Supongamos que Indiana decidió que el juego se había convertido en un problema severo para sus residentes: muchos de ellos se habían vuelto adictos a él y esta situación estaba abriendo paso a quiebras que sembraban la ruina en la vida familiar y en la economía de Indiana. Y, por consiguiente, decidió prohibir los casinos a lo largo del estado y plantear a los casinos de todos los demás estados la siguiente exigencia: si querían hacer negocios con

XXVII.7. Limitaciones a la jurisdicción personal por el debido proceso

Una compañía de Nueva York vende adminículos a un comprador de Oklahoma. La venta se negocia en Nueva York y los adminículos se embarcan allí, pero el vendedor tiene una pequeña oficina de ventas y algunas bodegas en Oklahoma. Si surge una disputa en los términos del contrato, ¿debiera permitirse que el comprador demandara al vendedor en Oklahoma o deberá ir a Nueva York a demandar?

Si se permite que el comprador demande en Oklahoma se impondrán costos adicionales al vendedor, no sólo de viaje sino también en el riesgo incrementado de perder el juicio si se instruye en un foro menos amistoso. Pero si el comprador debe demandar en Nueva York se le impondrá un aumento de costo similar. Un posible enfoque, que se utiliza para decidir cuál de dos o más tribunales con jurisdicción sobre una disputa debe ser el sitio efectivo del juicio, consiste en comparar los costos y requerir que el juicio se instruya en el estado donde se minimicen los costos totales de ambas partes. El problema de este enfoque (que lleva el nombre de *forum non conveniens*) es que otorga al demandante una gama de elección potencialmente muy grande de estados donde demandar, ¿pues quién sabe dónde se minimizarían realmente los costos de litigación conjuntos de las partes? La solución que los tribunales han adoptado consiste en requerir que el estado elegido por el demandante haya conferido por lo menos algún beneficio al demandado, de modo que el demandado no pueda quejarse demasiado si se ve obligado como un *quid pro quo* a defenderse en un foro que no es ideal desde su punto de vista. Esta condición se satisfaría en nuestro caso hipotético, dado que el demandado tiene cierta presencia tangible en Oklahoma.

No obstante, supongamos ahora que un vendedor de automóviles de Nueva York vende uno a un residente de esa ciudad que resulta lesionado, debido a un defecto del automóvil, mientras viaja a Oklahoma. ¿Podrá el comprador demandar al vendedor en Oklahoma? No podrá hacerlo.[17] En cierto sentido, el vendedor se beneficia de que Oklahoma tenga carreteras, así como la compañía telefónica de Nueva York se beneficia de que haya teléfonos en Oklahoma, pero el beneficio es demasiado pequeño.

Con todo, se podría engendrar un arbitraje en esta situación, en la que se compra a granel en un estado que no está sesgado a favor de los demandantes ubicados en él y se revende a un precio más elevado en un estado que esté sesgado a favor de sus demandantes locales. Los demandantes contarían con la ventaja del sesgo, pero el manufacturero-demandante no quedaría compensado al poder cobrar un cargo extra por sus artículos.

los residentes de Indiana, debían obtener una licencia de este estado que le prohibiría a cualquier habitante de Indiana apostar más de 10 dólares por día en un casino".

[17] Véase World-Wide Volkswagen Corp. *vs.* Woodson, 444 U.S., 286 (1980).

XXVII.8. Transferencias de agua entre cuencas

Dado que el agua es un bien excedente en algunos estados y un bien escaso en otros, parecería que hay buenas oportunidades para aumentar el valor total de ese recurso transfiriendo agua de las regiones superavitarias a las deficitarias. Es sustancial el costo de transportación del agua a largas distancias, pero los principales obstáculos para las transferencias interregionales son factores legales e institucionales derivados de la estructura del sistema federal.[18]

Si el comprador de fuera del estado pudiera obtener la cantidad de agua que necesita comprando derechos de apropiación individual poseídos por residentes de los estados que tienen agua en exceso, no se plantearía ningún problema especial, ya que un estado no podría prohibir la venta de un derecho de apropiación a un comprador de fuera del estado. Pero los grandes costos de transacción implicados en la compra de incluso un solo derecho de apropiación (véase antes III.11) volverían prohibitivos los costos del intento de agregar gran número de tales derechos por el comprador individual. Una alternativa es la compra de aguas no apropiadas, que se encuentran en grandes cantidades en algunos estados, sobre todo en el noroeste del Pacífico. Desafortunadamente, nadie puede transferir un título claro de tales aguas. Un derecho de propiedad sólo se crea por un acto de apropiación. Un no residente podría intentar una apropiación pero se enfrentaría al argumento de que las rentas de las aguas no apropiadas pertenecen a los residentes del estado donde se ubican. Y está implicado algo más que las rentas. Las aguas no apropiadas no son necesariamente aguas no usadas. Hay muchos usos del agua económicamente valiosos pero inapropiables, como la recreación y la dilución de contaminantes, que podrían verse perjudicados por una gran transferencia a otra cuenca hidráulica.[19]

El problema podría resolverse con el pago de una compensación apropiada del usuario de fuera del estado al estado de origen. Pero la repartición de la compensación entre los residentes afectados plantearía dificultades ad-

[18] Véanse Charles J. Meyers y Richard A. Posner, *Market Transfers of Water Rights: Toward an Improved Market in Water Resources*, parte 4 (informe de la Comisión Nacional del Agua, 1 de julio de 1971, publicado por el Servicio Nacional de Información Técnica); Rodney T. Smith, *Trading Water: An Economic and Legal Framework for Water Marketing* (1988); Stephen F. Williams, "Free Trade in Water Resources: Sporhase *vs.* Nebraska ex rel. Douglas", 2 *Supreme Ct. Econ. Rev.,* 89 (1983); George A. Gould, "Transfer of Water Rights", 29 *Natural Resources J.,* 457 (1989); Barton H. Thompson, Jr., "Institutional Perspectives on Water Policy and Markets", 81 *Cal. L. Rev.,* 671 (1993).

[19] Esto sugiere otra crítica del sistema de derechos de apropiación existente en el derecho del agua, además de la discutida antes en III.4: muchos usos valiosos del agua no poseen el carácter de asignación, de manera que en el caso de ellos no se pueden obtener los derechos de propiedad dentro del esquema de un sistema de derechos de apropiación y, como resultado, estos usos se encuentran menos desarrollados de lo que debieran estarlo.

ministrativas de índole sustancial, además de la ausencia, antes de la apropiación, de un propietario reconocido. Y cualquier título transferido por el estado estaría nublado por los intereses extensos pero incuantificados del gobierno federal y de otros estados. El gobierno federal posee derechos —cuya extensión no se ha determinado nunca— al uso de las aguas de los ríos del extenso dominio público federal en los estados occidentales (reservaciones indias, parques y bosques nacionales, bases militares, etc.). Tiene también intereses mal definidos pero extensos en las aguas de todos los ríos navegables en virtud de su servidumbre de navegación. Por lo tanto, una sustancial desviación entre cuencas requeriría la aprobación del gobierno federal, y es probable que sólo el Congreso pudiera otorgar esta aprobación. Requeriría también la aprobación de los otros estados de la cuenca del río, ya que la desviación de una cantidad de agua sustancial desde un punto del sistema del río puede afectar intereses recreativos, ambientales y otros intereses económicos valiosos a lo largo del sistema fluvial. Una definición de los intereses de cada estado requeriría una repartición por la Corte Suprema, un organismo interestatal o el Congreso.

XXVII.9. La exportación de la pobreza

El costo de los servicios gubernamentales que consumen los pobres desprovistos de trabajo es mayor a las recaudaciones fiscales que ellos aportan para la prestación de tales servicios: son un elemento de déficit en las finanzas públicas del estado. Si además tienen escaso poder político, como ocurre comúnmente (la participación de los pobres en las urnas es baja, y no hacen aportaciones a la campaña), habrá un fuerte apoyo para el intento de "exportarlos". El método más barato para lograr este cometido consiste en establecer un nivel de bienestar tan bajo que los beneficiarios se vean inducidos a emigrar a los estados que otorguen beneficios mayores.[20] Un problema similar afectó a la administración inglesa del auxilio a los pobres desde su inicio, lo que generó prohibiciones contra la salida de los pobres de su parroquia

[20] Otra forma consiste en entrar en tratos con el Departamento de Justicia y establecer "acuerdos relativos a la aplicación de las leyes de inmigración" merced a los cuales los funcionarios estatales o locales son, en efecto, nombrados sustitutos de los funcionarios de inmigración, quedando autorizados para investigar, aprehender y mantener en prisión a inmigrantes que pueden encontrarse en el país en forma ilegal. 8 U.S.C. 1357 (g) (1). Aparentemente, existen más probabilidades de que los estados provistos de beneficios de bienestar social más elevados suscriban tales acuerdos. Heather M. Creek y Stephen Yoder, "With a Little Help from Our Feds: Understanding State Immigration Enforcement Policy Adoption in American Federalism", 40 *Policy Stud. J.,* 674 (2012). Para un acopio de pruebas de que los inmigrantes ilegales les cuestan a los gobiernos estatales y locales más en términos de beneficios y servicios de lo que pagan de impuestos, véase Congressional Budget Office, The Impact of Unauthorized Immigrants on the Budgets of State and Local Governments, diciembre de 2007, www.cbo.gov/sites/default/files/cbofiles/ftpdocs/87xx/doc8711/12-6-immigration.pdf.

original. La solución tradicional de nuestros estados, una versión atenuada de la práctica inglesa, consistió en requerir que un individuo residiera en el estado durante cierto tiempo, un año de ordinario, antes de que pudiese recibir los beneficios de la beneficencia. A pesar de esta regla, muchos indigentes de los estados de beneficios bajos como Misisipi emigraron a los estados de beneficios altos como Nueva York, arreglándoselas a duras penas con el auxilio de parientes o amigos durante el año de espera. Pero no hay duda de que otros se vieron desalentados.

En Shapiro *vs*. Thompson,[21] la Corte Suprema invalidó los requerimientos de residencia para el derecho a beneficios de bienestar alegando que inhiben el viaje interestatal. Empero, la emigración inducida por la perspectiva de recibir pagos de bienestar o beneficencia mayores que los disponibles en el estado de origen del emigrante distorsiona, en lugar de promover, la distribución geográfica eficiente de la población. La sociedad no estará en mejor situación si los indigentes se mudan de Misisipi a Nueva York porque Nueva York tenga una beneficencia más generosa. Estará en una posición peor: los indigentes pueden ser apoyados decentemente a un costo menor en Misisipi que en Nueva York. No está siquiera claro que los indigentes como un grupo mejoren su situación con tal emigración, aunque algunos indigentes individuales sí lo hagan. La emigración libre alienta a los estados de baja beneficencia en su política de inducir la emigración mediante una asistencia pública inadecuada. Su efecto es más complejo en los estados de altos pagos de bienestar o beneficencia. Por una parte, el aumento del número de sus residentes indigentes podría aumentar la influencia política de los pobres en el estado, lo que se traduciría en niveles de apoyo mayores aún. Por otra parte, cada incremento atrae a nuevos receptores de beneficencia de otros estados y finalmente el estado deberá advertir que está permitiendo que los estados de beneficencia baja lo estén usando para aliviarse de la carga del sostenimiento de sus propios pobres. Una vez que el estado de alta beneficencia deja de aumentar su presupuesto destinado a ese objeto, disminuirán los pagos medios de beneficencia de sus residentes, porque los indigentes continuarán llegando, atraídos por un nivel de sostenimiento que, aunque estático y en promedio declinante (el pastel no está creciendo sino que se está dividiendo en más fragmentos), es todavía mayor que en sus propios estados.

El problema básico, tanto en la beneficencia inglesa como en la nuestra, es el de la organización regional, la que genera incentivos socialmente improductivos como el de los estados para transferir el costo de la beneficencia a otros estados y el de los pobres para emigrar hacia las áreas donde sean más generosos los pagos de la beneficencia. Un requerimiento de residencia es un método burdo y sólo parcialmente eficaz para abordar este problema: burdo porque a menudo puede desalentar la emigración de una familia pobre que

[21] 394 U.S., 618 (1969).

es atraída por las mejores oportunidades de empleo en el estado de altos pagos de bienestar o beneficencia pero necesita una asistencia pública moderada durante las primeras semanas o meses de residencia en un nuevo estado;[22] sólo parcialmente eficaz porque ese requerimiento invita al fraude y porque, como se señaló antes, una familia indigente podría ser capaz de sobrevivir durante un año sin la asistencia pública.

Dependiendo del equilibrio político existente dentro de un estado, algunos estados podrían desear atraer a personas pobres por la misma razón que el derecho de voto se ha expandido tan grandemente desde el siglo XVIII: la coalición política dominante estará a favor de la expansión de la franquicia, o de atraer votantes potenciales de otros estados o de países extranjeros, si esperan que los nuevos votantes apoyarán a la coalición. Algunos estados podrían incluso desear atraer a los que han quebrado, como lo lograrán con seguridad las exenciones más generosas contra las presiones de los acreedores; en virtud de que la quiebra le depara a un sujeto la oportunidad del borrón y cuenta nueva, estos deudores podrían ganar buenos ingresos y contribuir así a la base tributaria.[23] Pero, a pesar de estas circunstancias, es probable que Shapiro haya contribuido a la declinación de los pagos medios de la beneficencia en los últimos decenios, ya que, al volver a los estados de beneficencia generosa un punto de atracción para los pobres, aumentó los costos de la beneficencia generosa para los contribuyentes de los estados.

Si se nacionalizara la beneficencia, ¿sería una buena idea, desde un punto de vista económico, incluir un ajuste del costo de la vida, a fin de que los receptores en los estados de costo alto recibieran beneficios mayores que los de los estados de costo bajo? No. El ajuste haría que para los receptores pobres resultara indiferente vivir en áreas de un costo elevado o de un costo bajo. Pero nuestro deseo es que vivan en áreas de bajo costo a fin de minimizar los costos sociales derivados de mantenerlos. Un nivel uniforme de auxilio, independiente de las diferencias regionales en el costo de la vida, los alienta a decidirse por esta opción, puesto que un dólar ahorrado en costos de vida significa un dólar añadido a sus ingresos disponibles.

XXVII.10. REVISIONES DEL CONFLICTO DE LEYES Y DE LA CONTAMINACIÓN

El fallo de un tribunal estatal acerca de cuál será el estado cuya ley deberá aplicarse a cuestiones sustantivas en un juicio puede plantear una interrogante

[22] Nótese que el requerimiento de la residencia desalienta a los indigentes de los estados provistos de altos niveles de beneficios de bienestar social en lo tocante a la búsqueda de empleo en otros lugares, puesto que, si perdieran el trabajo en el otro estado, en el caso de que decidieran volver tendrían que afrontar el elevado costo del requisito de residencia.

[23] Véanse pruebas empíricas que apoyan ambas hipótesis en Margaret F. Brinig y F. H. Buckley, "The Market for Deadbeats", 25 *J. Legal Stud.*, 201 (1996).

en torno al federalismo. Consideremos dos estados, A y B. A alberga cierto número de grandes fabricantes de productos de consumo y, en consecuencia, su ley de responsabilidad del fabricante no tiene una inclinación marcada a favor del consumidor. B no alberga a tales fabricantes y su ley se inclina marcadamente a favor de los consumidores. Supongamos que debido a la posibilidad de arbitraje es inviable que los fabricantes vendan a precios mayores en B que en A a fin de reflejar los mayores costos de la responsabilidad en B: si trataran de hacerlo, los comerciantes de A comprarían para revender en B, ganándose la diferencia existente entre los precios. Esto significa que los consumidores de A se verán obligados a subsidiar a los consumidores de B.

Esto supone que los tribunales del estado B aplicarán su ley a un accidente sufrido por un residente de B en B, mientras que A aplicará su ley a un accidente sufrido por un residente de A en A. En efecto, éstos son los resultados según la regla tradicional del conflicto de leyes de *lex loci delicti* (la ley del estado donde ocurrió el accidente, lo que significa de ordinario el estado en el que el demandante fue lesionado). Pero ahora los estados, como vimos antes (en XXII.19), aplican a menudo de cara al conflicto de leyes un análisis más de "intereses" que está más libre de ataduras, según el cual los tribunales del estado A podrían aplicar la ley de C si C resultara ser el estado del fabricante cuyo producto perjudicó al residente de A y tuviera una ley de responsabilidad del fabricante más estricta que la de A. Todo estado productor de manufactura vacilará ante la perspectiva de adoptar una ley de responsabilidad del fabricante fuertemente inclinada en contra de los fabricantes.[24]

Las externalidades de la contaminación interestatal plantean problemas similares. Los contaminantes aéreos y acuáticos no respetan las fronteras estatales. Las emisiones de dióxido de azufre, por ejemplo, pueden provocar la lluvia ácida en estados distantes del estado de la emisión. Los vientos prevalecientes soplan de oeste a este, creando una diferencia sistemática entre los estados de donde viene el viento y los que están en dirección del viento y permitiendo que los primeros exporten mucho de la contaminación generada por las actividades industriales y otras emisoras de contaminación dentro del estado. En una división correcta de las responsabilidades entre las agencias federales y estatales de control de la contaminación, la Agencia de Protección Ambiental (EPA) de los Estados Unidos otorgaría una alta prioridad al control de las externalidades de la contaminación interestatal, porque no hay ninguna buena razón para suponer que los estados dejarán de proveer una protección óptima contra la contaminación para sus propios residentes.[25] Especialmente

[24] Véase Bruce L. Hay, "Conflicts of Law and State Competition in the Products Liability System", 80 *Geo. L. J.,* 617 (1992). No obstante, *cf.* Frank B. Cross, "Tort Law and the American Economy", 96 *Minn. L. Rev.,* 28 (2011).

[25] Véase Richard L. Revesz, "Rehabilitating Interstate Competition: Rethinking the 'Race-to-the-Bottom' Rationale for Federal Environmental Regulation", 67 *N.Y.U. L. Rev.,* 1210 (1992). ¿Podría un estado proteger *en exceso* a sus residentes? Sin duda alguna, precisamente de la for-

perversas son las reglas que, al prohibir una degradación considerable de la calidad del aire existente, limitan la capacidad de los estados que están en dirección del viento para desarrollarse industrialmente, asegurando así que el nuevo desarrollo industrial se concentre en los estados de donde viene el viento… y recompensándolos así por la contaminación que están causando. Pero sería un error responder haciendo a los contaminadores de donde viene el viento estrictamente responsables de la contaminación que causan en los que están en dirección del viento. La víctima de la contaminación podría ser quien la evita al costo menor (véanse antes III.7 y XIV.6). La formidable tarea de la EPA consiste en impedir, en la mayor medida posible, que la ubicación de una actividad contaminante se vea influida por las fronteras estatales.[26]

Los problemas creados cuando la víctima es quien puede evitar al menor costo una actividad contaminante se ilustran dramáticamente por los esfuerzos de algunos estados para impedir la importación de desperdicios tóxicos para su eliminación en rellenos sanitarios. La tarea para los tribunales que deben decidir si el esfuerzo del estado importador grava indebidamente al comercio interestatal consiste en determinar si los costos para el estado importador son más altos o más bajos que los beneficios *a escala nacional* de los rellenos sanitarios, ya sea que estos beneficios los reciba o no el estado.

XXVII.11. EL PODER DEL COMERCIO

Si se define al comercio interestatal en una forma suficientemente amplia, el poder del Congreso para regular dicho comercio abarca toda la actividad económica de la nación. Durante más de medio siglo se supuso que el poder del Congreso en ese respecto era en efecto pleno, hasta que la Corte Suprema, en United States *vs.* Lopez,[27] invalidó por no estar autorizada por la cláusula comercial una ley que convertía en un delito federal la portación de un arma de fuego (sin permiso legal para portarla) a menos de 300 metros de una escuela. La portación del arma de fuego no podría ser considerada como una transacción interestatal, pero el voto de disentimiento sostenía que, al interferir con el proceso educativo, la portación de armas de fuego en la vecindad de una escuela podría impedir la formación del capital humano de los estudiantes, reduciendo su productividad en los mercados nacionales. Esto es cierto, pero aun el efecto agregado de toda portación de armas de fuego a menos de 300 metros de las escuelas sobre la productividad de los trabajadores futuros debe de ser muy escaso. En la medida en que haya ventajas económicas para

ma en que procede el estado en favor del consumidor en el ejemplo de la responsabilidad del fabricante.

[26] Véase Richard L. Revesz, "Federalism and Interstate Environmental Externalities", 144 *U. Pa. L. Rev.,* 2341 (1996).

[27] 514 U.S., 549 (1995).

el federalismo (ventajas no sólo al evitar las deseconomías de escala y estimular la experimentación, sino también al aumentar la presión competitiva para que el gobierno actúe eficientemente), es esencial que se prohíba al gobierno federal arrebatar funciones estatales, lo que es un peligro si no se requiere que las actividades tengan efectos *sustanciales* sobre los mercados regionales o nacionales para que se permita la regulación federal.

Aquí es crítico el hincapié en los agregados. De otro modo, Lopez sería incongruente con Wickard *vs.* Filburn,[28] donde se sostuvo que la siembra de un cultivo por un agricultor se encontraba dentro del poder de comercio federal aunque ese individuo planeara consumir él mismo toda la cosecha en lugar de venderla. Por supuesto, la actividad de un agricultor no afectará perceptiblemente al comercio, pero, si ésa fuera la prueba, el gobierno federal no podría regular la agricultura, aunque sea uno de los mayores mercados interestatales y extranjeros de la nación. Muchos pocos hacen un mucho. Los economistas entienden esto bien, porque modelan la demanda de un bien o servicio producido en un mercado competitivo como la suma de curvas de demanda infinitamente elásticas que afrontan los vendedores individuales. Ninguno de ellos puede afectar el precio o la cantidad del mercado, pero el agregado de los vendedores sí puede hacerlo y lo hace.

El problema de Wickard *vs.* Filburn se planteó apropiadamente en términos de si el autoconsumo de los agricultores (y no de un agricultor) podría afectar considerablemente al mercado agrícola, y por supuesto la respuesta era que sí podría hacerlo. El precio y la cantidad, y por ende los embarques interestatales y extranjeros de un cultivo, digamos el maíz, se determinan por la demanda y la oferta, incluida la demanda de los propios cultivadores. Carece de importancia para el precio y la cantidad que los productores se coman una parte de su maíz sin venderlo, o vendan toda la cosecha de maíz y luego compren la cantidad de maíz que quieran comerse. La cantidad de maíz demandada y ofrecida es la misma en ambos casos, y esto demuestra que el consumo del productor es una parte del mercado nacional e internacional como lo son las ventas de maíz. Y es una parte lo suficientemente grande para afectar al precio, la producción y los embarques.

En la decisión de la Corte Suprema de defender la constitucionalidad de la Ley de Protección al Paciente y Cuidado Asequible (PPACA), hoy conocida como la Ley de Cuidado de la Salud a Bajo Precio (ACA), u Obamacare,[29] la mayoría de los magistrados sostuvieron para sorpresa de todos que la disposición de la ley que exige a las personas que no compren un seguro de salud el pago de una "penalización" (si es que les alcanza) —la provisión conocida como el "castigo"— sobrepasaba la autoridad del gobierno federal sobre el comercio interestatal. La Corte determinó que el castigo quedaba incluido

[28] 317 U.S., 111 (1942).

[29] National Federation of Independent Business *vs.* Sebelius, 132 S. Ct., 2566 (2012).

bajo el poder tributario del Congreso. Reconocía que la industria de seguros de salud de los Estados Unidos se halla "envuelta en" el comercio interestatal, pero juzgaba que el poder en materia comercial no autoriza al gobierno a obligar a las personas a involucrarse en el comercio interestatal. Ahora bien, la Ley no obligaba a nadie a comprar un seguro de salud; simplemente penalizaba el "gorroneo", esto es, el caso de las personas que aun cuando no cuentan con un seguro buscan tratamiento médico, lo reciben al fin (a menudo en salas de emergencia de los hospitales) y, al recibir la factura, resulta a menudo que no pueden pagarla, por lo que los costos son derivados a otros individuos. El castigo señalado en la Ley de Cuidado de la Salud a Bajo Precio responde al objetivo de internalizar ese costo externo y, así, lograr que la industria interestatal de los seguros de salud opere de manera más eficiente.

Colocar el estatuto bajo el poder tributario del Congreso era una decisión extraña porque, tal como se expresa en el Artículo 1, sección 8, cláusula 1 de la Constitución, que confiere ese poder, lo confiere con la finalidad de permitirle al gobierno federal financiarse a sí mismo. El mandato es un impuesto y habrá de generar cierta recaudación, pero la recaudación no constituye en sí su objetivo. Es un impuesto regulatorio (véanse antes XIV.6, XVII.3) encaminado a alterar el comportamiento. Entre más efectivo es ese impuesto, menos recaudación genera.

Lecturas recomendadas

Brinig, Margaret F. y F. H. Buckley, "The Market for Deadbeats", 25 *J. Legal Stud.,* 201 (1996).

Fiscal Federalism and the Taxation of Natural Resources (Charles E. McLure, Jr. y Peter Mieszkowski [comps.], 1983).

Hay, Bruce L., "Conflicts of Law and State Competition in the Products Liability System", 80 *Geo. L. J.,* 617 (1992).

Hellerstein, Walter, *State and Local Taxation of Natural Resources in the Federal System: Legal, Economic, and Political Perspectives,* caps. 4-6 (1985).

Levmore, Saul, "Interstate Exploitation and Judicial Intervention", 69 *Va. L. Rev.,* 563 (1983).

McLure, Charles E., Jr., "Incidence Analysis and the Supreme Court: An Examination of Four Cases from the 1980 Term", 1 *Supreme Ct. Econ. Rev.,* 69 (1982).

Posner, Richard A., *The Federal Courts: Challenge and Reform,* cap. 9 (1996).

Revesz, Richard L., "Federalism and Interstate Environmental Externalities", 144 *U. Pa. L. Rev.,* 2341 (1996).

Rose-Ackerman, Susan, "Does Federalism Matter? Political Choice in a Federal Republic", 89 *J. Pol. Econ.,* 152 (1981).

Shaviro, Daniel, "An Economic and Political Look at Federalism in Taxation", 90 *Mich. L. Rev.,* 895 (1992).

Siegel, Neil S., "Free Riding on Benevolence: Collective Action Federalism and the Minimum Coverage Provision", 75 *Law & Contemp. Prob.,* 29 (2012).
The Economics of Federalism (Bhajan S. Grewal, Geoffrey Brennan y Russell L. Mathews [comps.], 1980).
Voigt, Stefan y Lorenz Blume, "The Economic Effects of Federalism and Decentralization — A Cross-Country Assessment", 151 *Pub. Choice*, 229 (2012).
Weingast, Barry R., "The Economic Role of Political Institutions: Market Preserving Federalism and Economic Development", 11 *J. Law, Econ. & Organization,* 1 (1995).
Williams, Stephen F., "Severance Taxes and Federalism: The Role of the Supreme Court in Preserving a National Common Market for Energy Supplies", 53 *U. Colo. L. Rev.*, 281 (1982).

Problemas

1. ¿Por qué debiera abstenerse Montana, en nuestro ejemplo de un impuesto a la extracción, de gravar a los no residentes? ¿No se benefician los consumidores de carbón no residentes de los servicios brindados por Montana a los productores de carbón ubicados en el estado?

2. ¿Podrá la teoría económica ayudarle a encontrar una fórmula para la repartición de los bienes muebles de una empresa (por ejemplo, los aviones de una aerolínea) entre los estados para los fines del impuesto estatal?

3. ¿Debiera permitirse a los estados crear impuestos al ingreso corporativo? Supongamos que los estados que crean tales impuestos fuesen obligados a adoptar una fórmula uniforme para repartir el ingreso corporativo entre ellos. ¿Cuál sería una fórmula mejor: el porcentaje de los ingresos corporativos obtenidos de las ventas a residentes del estado que grava o el porcentaje de los empleados dentro del estado que grava? ¿Sería preferible el porcentaje de la propiedad tangible de la corporación ubicada en el estado?

4. En Baldwin *vs.* G.A.F. Seelig, Inc., 294 U.S., 511 (1935), sostuvo la Corte Suprema que una ley del estado de Nueva York que fijaba un precio mínimo para la leche vendida en el estado no podría aplicarse constitucionalmente a la leche producida en Vermont e importada a Nueva York. Se ha criticado como sigue la fuerte condena de la Corte contra el intento de Nueva York de someter esa leche a su ley del precio mínimo: "El interés de Nueva York no era un mero sesgo económico contra los competidores de fuera del estado, al ocuparse primordialmente del bienestar económico y la salud locales". Gerald Gunther y Noel T. Dowling, *Cases and Materials on Constitutional Law,* 651 (8ª ed., 1970).[30] ¿Puede diferenciarse el "sesgo económico contra los

[30] Ediciones subsecuentes han eliminado lo referente a la tributación estatal.

competidores de fuera del estado" del "bienestar económico local"? ¿Equivale esta crítica a la afirmación de que debiera permitirse a los estados la imposición de algunos aranceles?

5. ¿Por qué tienen algunos estados más prestaciones de bienestar que otros? ¿Es porque los contribuyentes de tales estados asignan un valor mayor a la disminución de la pobreza? En caso afirmativo, ¿podrá decirse todavía que la eficiencia disminuye, en lugar de aumentar, si los pobres se mudan a estos estados?

6. De acuerdo con la doctrina del "participante en el mercado" se permite que un estado discrimine a los vendedores no residentes si está participando efectivamente en el mercado, por ejemplo como un comprador, antes que regulando un mercado privado. Así, un estado podría tener una regla de comprar clips de papel sólo a vendedores residentes aunque no pudiera obligar a los compradores privados de clips a adoptar esa política. ¿Tendría esta doctrina buen sentido económico si los estados tuvieran los mismos incentivos económicos que los particulares? ¿Si un estado quisiera transferir riqueza de los no residentes a los residentes? ¿Podría haber tal transferencia si el mercado del producto en cuestión fuera nacional antes que estatal o local? Véase W. C. M. Window Co. *vs*. Bernardi, 730 F2d, 486 (7th Cir., 1984).

7. Según las interpretaciones actuales de la ley antimonopólica federal, las agencias estatales son inmunes a la responsabilidad, pero no los municipios y otras agencias locales, a menos que actúen bajo la dirección del gobierno estatal. ¿Es congruente este patrón con la economía del federalismo? ¿Sería quizá más congruente el patrón opuesto? Véase Frank H. Easterbrook, "Antitrust and the Economics of Federalism", 26 *J. Law & Econ*., 23 (1983).

8. Los estados compiten por los negocios ofreciendo concesiones fiscales, mejoras de la infraestructura y otros beneficios monetarios o monetizables. ¿Ilustra tal competencia los beneficios sociales, o los costos sociales, del federalismo? Véase Peter D. Enrich, "Saving the States from Themselves: Commerce Clause Constraints on State Tax Incentives for Business", 110 *Harv. L. Rev.,* 377 (1996).

9. En Collins *vs*. Heritage Wine Cellars, Ltd., 589 F.3d, 895 (7th Cir., 2009), los demandantes eran camioneros que transportaban vino de un almacén en el área de Chicago, una fábrica propiedad del demandante Heritage, a tiendas minoristas en Chicago y otras áreas de Illinois. Con el propósito de acarrear el vino al almacén desde los estados y los países en los que se producía (en ninguna parte de Illinois se fabricaba vino), Heritage hizo contratos con compañías de camiones y otros transportistas. Se trataba de contratistas independientes. Ni ellos ni sus subordinados eran empleados de Heritage, a diferencia de los demandantes. No obstante, Heritage controlaba la operación de tránsito del vino de un lugar a otro y dirigía los movimientos a lo largo de toda la travesía desde el estado o país de origen hasta las tiendas de menudeo de Illinois adonde lo llevaban los demandantes desde las bodegas. Alrededor de

una cuarta parte del vino que Heritage embarcaba a su almacén en Illinois ya la habían solicitado por anticipado los dueños de las tiendas de menudeo. En virtud de que Heritage ya tenía en la mano una orden levantada, el vino permanecía en el almacén sólo por breve tiempo. Las tres cuartas partes restantes del líquido permanecían guardadas en el almacén en espera de que Heritage lograra encontrar un comprador; la mayor parte de ese vino sufría rotación cada mes, tras haber sido comprado y embarcado por Heritage sobre la base de sus estimaciones acerca de la demanda de los clientes. Ninguna porción del vino estaba sujeta a alteración alguna en su viaje desde los viñedos a una tienda al menudeo en Illinois. No existía ningún procesamiento (ni siquiera un esfuerzo deliberado de contribuir a añejar el vino en el almacén), no se añadían aditivos, no se incorporaba el producto a otro producto, ni siquiera se lo reetiquetaba como una marca privada (marca de la casa). Cuando los contenedores de vino arribaban al almacén, los sacaban de los jergones de fundas retráctiles en que venían empaquetados y los colocaban en los estantes de las bodegas.

No obstante, ¿está la parte del transporte que se realiza totalmente dentro de Illinois inscrita "en" el comercio interestatal (en lugar de ser parte de aquello que afecta al comercio interestatal)?

XXVIII. LA DISCRIMINACIÓN RACIAL

XXVIII.1. El gusto por la discriminación[1]

A algunas personas no les gusta asociarse (salvo en una relación de supervisión o control: una excepción importante[2]) con miembros de grupos raciales, religiosos o étnicos diferentes del suyo, y pagarán un precio por darse ese gusto. Por ejemplo, hay ganancias pecuniarias en el comercio entre negros y blancos —que los negros trabajen para blancos (o viceversa), que los blancos vendan casas a negros, etc.—, así como hay ganancias pecuniarias en el comercio entre naciones. Sin embargo, al aumentar el contacto entre los miembros de las dos razas, tal comercio impone costos no pecuniarios, pero reales, a los miembros de cualquiera de las razas que no gustan de asociarse con miembros de la otra. Estos costos son análogos a los costos de transportación en el comercio internacional, que también reducen la cantidad del comercio.

No hay nada ineficiente en esto, pero los efectos de riqueza pueden ser dramáticos. Supongamos que a los blancos no les gusta asociarse con negros, pero los negros se sienten indiferentes ante la identidad racial de aquellos con quienes se asocian. Los ingresos de muchos blancos serán menores de lo que serían si no tuvieran ese gusto. Esos blancos renuncian a intercambios ventajosos; por ejemplo, podrían negarse a vender sus casas a negros que estén dispuestos a pagar precios mayores a los de los compradores blancos. Los blancos no prejuiciados gozarán de una ventaja competitiva al tratar con los miembros de la minoría poco gustada, al igual que estar librado del miedo a las alturas brinda una ventaja en la competencia por un empleo que implique elevarse a las alturas. Si nadie fuera desafecto a los negros o a las alturas, nadie podría disponer de una prima por trabajar conviviendo con negros o subiéndose al punto cumbre de las edificaciones.

No obstante, la antipatía racial de los blancos prejuiciados hará disminuir no solamente sus ingresos sino asimismo los ingresos de los negros, al impedirles a éstos realizar intercambios ventajosos con ellos; y la disminución de los ingresos de los negros será proporcionalmente mayor a la disminución de los ingresos de los blancos. Porque los negros son sólo una parte pequeña de

[1] Véase Gary S. Becker, *The Economics of Discrimination* (2ª ed., 1971), y, para una confirmación empírica de gran parte (aun cuando no la totalidad) de la teoría planteada por Becker, Kerwin Kofi Charles y Jonathan Guryan, "Prejudice and Wages: An Empirical Assessment of Becker's *The Economics of Discrimination*", 116 *J. Pol. Econ.,* 773 (2008).

[2] Stewart J. Schwab, "Employment Discrimination", en *Labor and Employment Law and Economics,* 296 (Kenneth G. Dau-Schmidt, Seth D. Harris y Orly Lobel [comps.], 2009).

la economía, el número de intercambios ventajosos que los negros pueden hacer con blancos es mayor al número de transacciones ventajosas que los blancos pueden hacer con negros. El sector blanco es tan grande que resulta virtualmente autosuficiente; el sector negro es mucho menor y más dependiente del comercio con los blancos.

La analogía del comercio internacional puede ayudar a aclarar el punto. Los Estados Unidos constituyen un agregado tan grande de habilidades, recursos y población que podrían sobrevivir a una disminución sustancial de su comercio exterior con relativa comodidad. Suiza no podría hacerlo. Sus propios mercados son demasiado pequeños y sus recursos demasiado limitados para permitirle alcanzar las economías de escala y de especialización sin comerciar con otros países. La posición de la minoría negra en los Estados Unidos es similar a la de Suiza en la economía mundial.

Aunque la discriminación es congruente con la competencia, así como una disminución del comercio internacional debida a los mayores costos de transportación no sería una prueba de que los mercados internacionales no son competitivos, en los mercados competitivos operan constantemente fuerzas económicas que tienden a minimizar la discriminación. En un mercado de muchos vendedores variará la intensidad del prejuicio contra los negros. Algunos vendedores tendrán sólo un prejuicio leve contra ellos y, por ello, no sacrificarán tantas transacciones ventajosas con los negros como sus competidores más prejuiciados (a menos que interfiera la ley).[3] Por lo tanto, sus costos serán menores, y esto les permitirá aumentar su participación en el mercado. Con todo, si bien los vendedores menos prejuiciados pueden llegar a dominar el mercado de la misma forma que las personas menos temerosas de las alturas acaban por dominar los empleos que demandan subirse hasta arriba, éste no es un desenlace inevitable. Un cierto grado de discriminación es eficiente (véase antes XII.10, y más adelante XXVIII.5), de modo que persistirá independientemente de que los propietarios o administradores de una empresa tengan algún gusto por la discriminación. Lo mismo ocurrirá (o podría ocurrir) con la discriminación que refleje los gustos de los consumidores antes que de los vendedores, pues aquéllos no afrontan presiones competitivas para cambiar sus gustos.

Y supongamos ahora que el mercado es monopólico en vez de competitivo. Es de esperarse que el vendedor único en el mercado sea, en promedio, tan prejuiciado como el miembro promedio, no menos prejuiciado, de la comunidad. Es cierto que existe la probabilidad de que cualesquiera monopolios que puedan transferirse libremente (como las patentes) tiendan a caer en las manos de los menos prejuiciados. Un monopolio que requiera la asociación con los negros es menos valioso para un propietario prejuiciado; éste sufre

[3] Véanse pruebas de que los trabajadores negros se benefician de la competencia entre los empleadores en Price V. Fishback, "Can Competition among Employers Reduce Governmental Discrimination? Coal Companies and Segregated Schools in West Virginia in the Early 1900s", 32 *J. Law & Econ.,* 311 (1989).

una disminución de su ingreso pecuniario al renunciar a transacciones ventajosas con negros, o un costo no pecuniario al realizar tales transacciones. Por lo tanto, el menos prejuiciado tenderá a comprar monopolios a los más prejuiciados. Pero no todos los monopolios son libremente transferibles.

Si el monopolio está regulado, las fuerzas del mercado que operan contra la discriminación se verán más debilitadas aún. Un procedimiento para evadir un tope a las ganancias consiste en sustituir el ingreso pecuniario por el no pecuniario, ya que el pecunario es muy difícil de controlar por una agencia reguladora; y un tipo de ingreso no pecuniario es la libertad para no asociarse con las personas contra quienes estamos prejuiciados.[4]

Los sindicatos que tienen poder monopólico podrían disminuir la eficacia de la competencia al nivel del vendedor para minimizar la discriminación. Un sindicato monopólico que aumente los salarios por encima del nivel competitivo generará un exceso de demanda de los empleos en los que se pagan estos salarios. Si el sindicato controla los empleos, tendrá que repartirlos de algún modo. Podría subastar las vacantes a medida que se presenten, o permitir que los miembros vendan su participación sindical; o podría adoptar criterios distintos de los precios, como el nepotismo o, como antes lo hacían los sindicatos, la pertenencia a la raza blanca. Los miembros del sindicato tomaban una parte de sus ganancias monopólicas bajo la forma de la liberación de un tipo de asociación que les disgustaba.[5]

Así pues, la política gubernamental, que es responsable de controlar las ganancias de los monopolistas y de crear sindicatos fuertes, podría aumentar la discriminación por encima del nivel que existiría en un mercado no regulado. Y éstas no son las únicas políticas gubernamentales que tienen un efecto adverso sobre las minorías raciales. Otra es la del salario mínimo. Los trabajadores que son objeto de discriminación deben aceptar un salario más bajo para superar la aversión de sus empleadores a contratarlos. Este proceso se ve impedido por la aplicación de un salario mínimo uniforme.[6]

XXVIII.2. La segregación escolar

En Brown *vs*. el Departamento de Educación,[7] la Corte Suprema invalidó las leyes estatales que exigían o permitían la segregación racial de las escuelas

[4] Véanse pruebas en Armen A. Alchian y Reuben A. Kessel, "Competition, Monopoly, and the Pursuit of Money", en *Aspects of Labor Economics,* 157 (Nat'l Bur. Econ. Research, 1962).

[5] Otra explicación, sugerida antes, es que la raza es un método barato para el racionamiento del acceso y para aumentar así las ganancias netas del monopolio de la oferta de mano de obra. Véase antes XII.10. Las dos explicaciones tienen las mismas consecuencias para el bienestar de los negros excluidos.

[6] Véanse antes XXII.8; Harold Demsetz, "Minorities in the Market Place", 43 *N.C. L. Rev.,* 271 (1965).

[7] 347 U.S., 483 (1954).

públicas. La Corte sostuvo que la educación segregada era inherentemente desigual porque instilaba un sentimiento de inferioridad en los niños negros. El análisis de la sección anterior sugiere una base económica, por oposición a una psicológica, para rechazar la noción de lo separado pero igual. La segregación disminuye las oportunidades de asociaciones valiosas entre las razas, y estas asociaciones serían especialmente valiosas para los negros a causa de la posición dominante de los blancos en la sociedad. La Corte había reconocido este punto en Sweatt *vs.* Painter,[8] donde se sostuvo que los negros no podían ser excluidos de las escuelas de derecho estatales. Señaló la Corte que los estudiantes negros de una escuela de derecho segregada no tendrían oportunidad de desarrollar valiosos contactos profesionales con los estudiantes que tenían mayores probabilidades de ocupar posiciones importantes en la judicatura y la barra después de la graduación. Rechazó los argumentos de que esta desventaja se contrarrestaba por la desventaja de los estudiantes blancos al verse privados de una asociación con los estudiantes de derecho negros, señalando que la débil posición de los negros en la profesión volvía tales asociaciones menos valiosas para los estudiantes blancos.

Si es correcto nuestro análisis anterior, las leyes invalidadas en Brown volvían a la discriminación mayor de lo que habría sido en ausencia de tales leyes… aunque quizá no mucho mayor. Si bien los tribunales federales, el Departamento de Justicia y otras agencias pudieron obligar a los estados sureños a dejar de aplicar las leyes de segregación, muchos blancos estaban dispuestos a pagar los costos adicionales necesarios para perpetuar la segregación escolar. Enviaron a sus hijos a escuelas privadas segregadas o se mudaron a distritos escolares que albergaban a pocos residentes negros. La Corte Suprema había vuelto la discriminación más costosa, pero, dado que la población blanca valuaba tal discriminación en alta medida, el efecto de Brown sobre la magnitud de la discriminación fue escaso durante muchos años. Además, dado que la población blanca controlaba las finanzas públicas de los estados, podía desviar la fuerza de la acción de la Corte subsidiando la educación privada mediante becas y créditos fiscales, medidas con las que se volvía más barato que los padres cambiaran a sus hijos a escuelas privadas segregadas.

El análisis económico podría ayudar a diseñar los decretos de desegregación, algunos de los cuales aún están siendo aplicados e impugnados. Supongamos que un tribunal desea promover la integración de las escuelas públicas de una comunidad que practicaban la segregación en el pasado (de modo que pueden ser colocadas bajo un decreto remedial), sin causar tanta "fuga de blancos" que los niños negros no obtengan ningún beneficio del decreto. Desde el punto de vista de los padres blancos que por cualquier razón consideran la presencia de niños negros como un detrimento para sus pro-

[8] 339 U.S., 629 (1950).

pios hijos, todo decreto de desegregación operará como un impuesto. Entre mayor sea el impuesto, más probable será que los padres blancos incurran en los costos de la mudanza a otro distrito escolar o del envío de sus hijos a la escuela privada. El tribunal puede minimizar este efecto (y, por ende, maximizar el beneficio del decreto para los negros) a: *1)* impartir al decreto de un alcance geográfico tan amplio como sea posible, de modo que se maximicen los costos de la reubicación para las familias blancas; *2)* imponer a los negros, y no a los blancos, el mayor número de los costos que sea posible, como cuando se transporta a la escuela a los niños negros antes que a los blancos, y *3)* limitar la fracción de cualquier escuela que es negra, ya que el "impuesto" de la desegregación sobre los blancos aumenta con la razón de niños negros a blancos en la escuela.

Aun si los niños negros se benefician grandemente de la educación integrada, no significa que no podrían beneficiarse más aún de otras estrategias. Por ejemplo, la Corte Suprema en Brown, en lugar de invalidar la segregación de la escuela pública, podría haber aprovechado el valor que otorgan a la segregación los blancos sureños exigiendo, como una condición de mantener las escuelas segregadas, que los estados sureños destinaran sumas mucho mayores que antes a la educación de los negros. Es concebible que los negros estuviesen mejor bajo tal arreglo aunque la decisión de Brown se hubiese cumplido pronto y a fondo. Imaginemos una comunidad integrada por 200 negros y 800 blancos, donde el ingreso medio de los negros es de 5 000 dólares y el de los blancos es de 10 000. Supongamos que la eliminación de la educación segregada aumentaría el ingreso pecuniario y no pecuniario de los negros en un promedio de 2 000 dólares (olvidemos la brecha existente entre el cambio de las condiciones educativas y el mejoramiento del empleo). Por lo tanto, la comunidad negra ganaría 400 000 dólares con la desegregación. Pero supongamos que los blancos de la comunidad estarían dispuestos a pagar un promedio de 1 000 dólares por persona para que no se integraran las escuelas. Por lo tanto, estarían dispuestos a pagar 800 000 dólares en el mejoramiento de la educación como el precio cobrado porque siga la segregación, y supongamos que cada dólar así gastado beneficiaría a los negros por un dólar. Entonces este gasto aumentaría los ingresos de los negros en 400 000 dólares más que la integración.

Esta estrategia alternativa no funcionaría para todas las instalaciones públicas segregadas. Los negros no pueden ser compensados por el insulto implícito en un régimen de baños y fuentes de sodas racialmente segregados por un decreto judicial que obligue al estado a gastar tanto dinero en las instalaciones de los negros como en las de los blancos. Por añadidura, si Brown hubiese meramente modificado, en lugar de invalidado, la segregación escolar pública, la decisión no habría tenido el efecto galvanizador que tuvo sobre el movimiento de derechos civiles.

XXVIII.3. Los requerimientos de la acción estatal

La Decimocuarta Enmienda, promulgada primordialmente para beneficio de las minorías raciales, establece que ningún *estado* negará a nadie la protección igual de sus leyes, ni privará a nadie de la vida, la libertad o la propiedad sin el debido proceso legal. El análisis económico puede ayudar a aclarar los problemas implicados en distinguir entre la acción estatal y la privada.

Pueden distinguirse tres niveles de la participación estatal en la discriminación: una ley u otra acción oficial que ordena la discriminación; la discriminación realizada por una empresa pública, y la participación estatal en empresas privadas que practican la discriminación pero no en la decisión de la empresa de discriminar. Los dos primeros niveles de la participación estatal estuvieron implicados en el caso de Brown, pero no se distinguieron. Podría pensarse que las leyes que exigen la segregación de todas las escuelas públicas de un estado promulgan los prejuicios de la mitad más prejuiciada de la población, de modo que generan una discriminación mayor que la habida si la decisión de segregar se dejara a los distritos de escuelas públicas individuales. No obstante, la Corte invalidó también las leyes estatales que permitían a los distritos escolares locales segregar a su elección, y cuando la decisión de segregar se deja a cada distrito escolar local no es tan obvio que el resultado será una cantidad de discriminación mayor a la que habría si toda la educación fuera privada; pero es probable que haya más. Un sistema de escuelas públicas es un monopolio no transferible (la educación privada, puesto que le cuesta al consumidor, diferenciado del contribuyente, más que la educación pública, no es un buen sustituto de esta última para privar a un distrito escolar público de todo su poder monopólico), y vimos antes que es de esperarse que los monopolios no transferibles discriminen más, en promedio, que las empresas competitivas o los monopolios libremente transferibles. Dado que la mayoría de los servicios gubernamentales tienen la naturaleza de monopolios no transferibles, esta observación tiene una aplicación general a las dependencias públicas.

El análisis es diferente cuando la decisión de discriminar la toman un individuo o una empresa privada, aunque el estado esté implicado en alguna medida en la actividad privada. Lo que debe determinarse es si la participación estatal vuelve más probable la discriminación. Cuando esa participación asume la forma de la regulación de un servicio público o un transporte público, como vimos antes, es en efecto mayor la probabilidad de que la empresa discrimine. El estado mantiene también un extenso sistema de registro de títulos de la tierra y participa de otro modo profundamente en la regulación del uso de la tierra. Pero esto no aumenta la probabilidad de que un propietario blanco de una casa se niegue a venderla a un comprador negro porque le disgusta asociarse con negros.

El análisis precedente sugiere una definición de la acción estatal diferente de la que han empleado los tribunales. Tal acción prohibiría la discriminación racial de los sindicatos, ya que las políticas gubernamentales que han promovido el crecimiento de sindicatos monopólicos han aumentado así la probabilidad de que practiquen la discriminación racial. Pero no prohibiría la discriminación del concesionario privado de un edificio público de oficinas,[9] a menos que la autoridad pública hubiera alentado al concesionario para que discriminara.

Supongamos que la participación estatal asume la forma del cumplimiento legal de una decisión privada de discriminar. ¿Podrán ejecutarse los convenios raciales?[10] ¿Podrá la ciudad de Macon, como administradora del parque donado por el senador Bacon, cumplir la condición racial de la donación?[11] ¿Prohíbe la cláusula de la protección igual que los comerciantes que no desean clientes negros recurran a los remedios civiles y penales contra el allanamiento? Resulta difícil creer que sin los derechos de propiedad habría una discriminación menor.

Es cierto, pero trivial, que si el Estado hiciera cumplir todas las decisiones privadas, excepto las que discriminen, el costo de la discriminación sería mayor y la incidencia menor. Una observación más interesante es que el efecto del cumplimiento de una condición racial en los casos del convenio y el fideicomiso de caridad restrictivos sería la creación de una discriminación mayor de la que desearían los miembros de la sociedad en la actualidad. Volviendo a la analogía del comercio internacional, esto es un poco como si las naciones hubiesen acordado en el siglo XIX que jamás permitirían que el comercio se realizara en algo que no fuesen barcos de vela. Éste es un ejemplo de la preocupación más amplia discutida en el capítulo XIX de que una condición perpetua en una escritura de propiedad o una donación pudiera hacer que los recursos se emplearan ineficientemente si se presenta una contingencia, en este caso una disminución del gusto por la discriminación. Pero es fortuito que el resultado de una condición perpetua sea una mayor o menor discriminación que la deseada por los contemporáneos. Si hubiera un aumento secular de la discriminación racial, antes que una disminución, el cumplimiento de las restricciones de la escritura o la donación racialmente motivadas (como una cláusula de la escritura constitutiva de una fundación que declare que su propósito es la promoción de la integración racial) podría generar una discriminación menor de la deseada por los contemporáneos.

[9] Sin embargo, véase Burton *vs.* Wilmington Parking Authority, 365 U.S., 715 (1961).
[10] Shelley *vs.* Kraemer, 334 U.S., 1 (1948).
[11] Véase Evans *vs.* Newton, 382 U.S., 296 (1966); antes el 19.3.

XXVIII.4. Las leyes contra la discriminación

Las leyes federales que prohíben la discriminación privada en la venta y la renta de inmuebles, en el empleo y en los restaurantes, hoteles y otros lugares de reunión pública tratan a menudo de justificarse primero como necesarias para eliminar los efectos de siglos de discriminación y luego como una promoción del comercio interestatal. La segunda justificación parece artificial para muchos observadores, pero tiene sentido económico. La discriminación disminuye las transacciones entre negros y blancos, y muchas de las transacciones que así se impiden se realizarían en el comercio interestatal, aun si éste se define estrictamente. La primera justificación es verosímil pero indefinida. *Toda* privación que padezcan ahora los negros podría deberse en parte a la discriminación del pasado. Si los niños negros se desempeñan en promedio menos bien que los blancos, incluso en las escuelas del norte, ello podría deberse a que el rendimiento financiero de la educación de los negros ha sido tradicionalmente bajo debido a una discriminación severa en el empleo contra los negros educados. Esta clase de argumento provee la justificación más fuerte para revertir la discriminación.

El análisis económico ayuda a explicar la variación en el cumplimiento de las leyes antidiscriminatorias. Si las asociaciones interraciales generadas por tal ley son escasas, será bajo el costo de la asociación aun para las personas prejuiciadas, quienes no estarán dispuestas a incurrir en grandes costos bajo la forma de un castigo, o de los costos legales, por la resistencia a cumplir la ley a fin de satisfacer su gusto. Así, no es sorprendente que haya habido un cumplimiento general de las leyes que prohíben a la gente negarse a vender inmuebles por razones raciales, aunque se hayan destinado escasos recursos al cumplimiento forzoso de estas leyes. A menos que el vendedor planee permanecer en la vecindad, su asociación con un comprador negro se limita a la negociación de la venta (y de todos modos lo hace un corredor). Por lo tanto, la mayoría de los casos de discriminación en la vivienda implican a las rentas más bien que a las ventas. La asociación entre el dueño de un hotel y su personal por una parte y los huéspedes del hotel por la otra es impersonal, excepto cuando el establecimiento es muy pequeño —y por esta razón los establecimientos pequeños fueron eximidos de la ley de habitación pública—, de modo que de nuevo no es sorprendente que el cumplimiento generalizado se haya logrado con rapidez y facilidad. La integración escolar es diferente. No sólo es la asociación entre los escolares íntima y prolongada, sino que en la medida en que los niños negros, en promedio, se desempeñan en la escuela peor que los niños blancos, la integración podría implicar para los blancos costos adicionales a los costos no pecuniarios impuestos por una asociación que no se desea.

Las leyes que prohíben la discriminación en el empleo implican interesantes cuestiones de prueba, de propósito estatutario, de remedio y de eficacia. Una

empresa podría no tener empleados negros, aunque se ubique en un área de gran población negra, por razones no relacionadas con la discriminación por parte de los administradores de la empresa o de los trabajadores blancos. Es posible que no haya negros con la capacitación o la aptitud requeridas, o que a los negros no les guste el tipo de trabajo, o que simplemente ignoren las oportunidades de empleo existentes en la empresa. Si un empleador se ve obligado a contratar negros poco calificados, pagarles una prima para inducirlos a hacer un tipo de trabajo que no les gusta o anunciar en la comunidad negra las vacantes de empleos que interesan a muy pocos negros, la empresa incurrirá en costos mayores a los beneficios para los negros que son contratados; el empleado negro no calificado impone pérdidas de productividad que él no recupera bajo la forma de salarios mayores; y la prima pagada al empleado negro a quien no le gusta trabajar en este tipo de empleo es un costo para la empresa pero no un beneficio para el empleado negro; simplemente compensa el costo no pecuniario del empleo para él. Es posible que el anuncio de empleos vacantes en la comunidad negra no confiera un beneficio conmensurable con sus costos si tal anuncio no genera un flujo considerable de solicitantes calificados. Dado que la mayoría de los costos adicionales se trasladarán probablemente a los clientes de la empresa, estos métodos de mejoramiento del bienestar de los negros son regresivos además de ineficientes.

Así pues, si los negros reciben un trato peor (menos pago y más desempleo) que los blancos debido a su manejo deficiente de habilidades y no a la propensión de los blancos a discriminarlos, es improbable que las leyes de antidiscriminación les presten auxilio, y el enfoque debiera desplazarse al empeño en mejorar sus habilidades.[12]

Las leyes que prohíben la discriminación en el empleo (véase también antes XII.10) son costosas aun cuando se apliquen a empleadores que efectivamente discriminan. Es posible que el empleador deba pagar un salario mayor a los trabajadores blancos que gusten de la discriminación y tengan alternativamente oportunidades de empleo atractivas en empresas que no cuenten con empleados negros. Si carecen de tales oportunidades, la eliminación de la discriminación podría no imponer costos pecuniarios —por hipótesis, los trabajadores no tienen más remedio que aceptar asociarse con negros—, pero impondrá costos no pecuniarios bajo la forma de una asociación que disgusta a los blancos. Y es improbable que los costos se vean contrarrestados por las ganancias de los trabajadores negros para quienes los empleos de la empresa son superiores a sus oportunidades de empleo alternativo o por las ventajas económicas que el aumento del comercio con los negros genera para la empresa y, por ende, para sus clientes; pues, si hubiera tales ganancias compensatorias, los negros habrían sido contratados sin ninguna presión legal.

[12] Pedro Carneiro, James J. Heckman y Dimitriy V. Masterov, "Labor Market Discrimination and Racial Differences in Premarket Factors", 48 *J. Law & Econ.,* 1 (2005).

Adviértase asimismo que los negros pagan como consumidores y como trabajadores su parte proporcional de todos los costos que impongan a las empresas las leyes contra la discriminación; pero comparten estos costos con los blancos, mientras que los beneficios sólo llegan a ellos (los negros). Además, no obstante, entre más cueste a las empresas el empleo de trabajadores negros, mayores serán los esfuerzos que hagan para minimizar su empleo de negros. Por ejemplo, se sentirán menos inclinadas a ubicar sus plantas u oficinas en áreas de densa población negra, sobre todo si, como se supone en la teoría del impacto dispar de la discriminación, las empresas serán más vulnerables a las acusaciones de discriminación entre mayor sea la población negra en el área en la que se ubican sus plantas y oficinas. Por añadidura, los empleadores pueden temer que si contratan a un negro y éste demuestra ser un trabajador poco satisfactorio resultará costoso despedirlo, pues probablemente los demande y, si gana el juicio, su pago de daños casi con toda certeza será bastante más cuantioso que si los demandara por no haber sido contratado y, por consiguiente, haber perdido una oportunidad de empleo en lugar de un empleo.

¿Cuál debería ser el remedio en un caso en el que se falla que un empleador ha discriminado? El análisis económico sugiere que deberá obligarse al empleador a pagar los daños de la persona discriminada, quizá duplicados o triplicados para facilitar la ejecución en los casos en que los daños son pequeños. Esto tendrá a la vez un efecto de compensación y de disuasión y parece preferible a un remedio que requiera que el empleador contrate un número especificado o un porcentaje de negros. Esa orden obligará al empleador a despedir trabajadores blancos o, lo que equivale a lo mismo, favorecer a los solicitantes negros sobre los blancos hasta que se alcance la cuota fijada en el fallo. Al imponer costos a los empleados blancos que podrían no tener la mácula de ser discriminadores a fin de mejorar la condición de los trabajadores negros, tal orden opera como un impuesto caprichoso y regresivo sobre la clase trabajadora blanca.

El análisis es mucho más complicado si los empleados comparten con el empleador la responsabilidad por la discriminación. Es posible que los empleados hayan excluido a los negros de su sindicato. O que el empleador haya discriminado únicamente porque sus trabajadores gustan de la discriminación, mientras que él mismo está libre de ella. De hecho, es más probable que un trabajador esté embargado de una animosidad discriminatoria a que lo esté un empleador. Trabaja en proximidad estrecha con los trabajadores negros, de suerte que si no le gustan los negros será más factible que incurra en una desutilidad mayor de la que distinguiría al remoto empleador.

Supongamos que un empleador les paga más a los trabajadores blancos que a los trabajadores negros en puestos de la misma clasificación sin ninguna justificación (como una brecha entre las habilidades de los trabajadores

blancos y las de los negros a las órdenes del empleador).[13] ¿Debiera ser la medida de los daños la diferencia existente entre las dos tasas salariales? Ése es el enfoque que adopta el derecho, pero sería necesario un análisis más complicado para determinar la pérdida real de los negros. Si el empleador se ve empujado a elevar el salario de los negros, sus costos marginales serán más altos, lo que desembocará en una contracción de la producción y, por ende, en despidos, de manera que muchos empleados negros podrían no salir beneficiados de una eliminación de la brecha entre los salarios de los blancos y los que perciben ellos. Y, si sólo se da el caso de que algunos empleadores en el mercado discriminan, no habrá ninguna repercusión sobre los salarios de los negros; los trabajadores de color simplemente laborarán para otros empleadores y serán remplazados por los blancos.

Podríamos suponer que el número de los juicios contra la discriminación disminuiría con el tiempo, a medida que disminuyera el prejuicio entre los vendedores en mercados competitivos. En realidad, ese número ha aumentado. Esto no refuta la teoría económica de la discriminación. A medida que los empleadores blancos contratan a más y más negros, la composición de las reclamaciones contra la discriminación cambia de las negativas de contratación a los despidos. Los juicios por despido son más lucrativos para los demandantes porque el monto de los daños se basa en los sueldos de mediados de la carrera antes que en los sueldos iniciales que son de ordinario menores, y en virtud de que los primeros tienen mayores probabilidades de exceder el costo de oportunidad —ya sea para compensar la inversión del empleado en el capital humano específico de la empresa o para crear un castigo adicional por eludir sus obligaciones (véase antes XII.10)—. Además, hay más dimensiones a lo largo de las cuales un empleador puede discriminar tanto a un empleado como a un solicitante (por ejemplo, el acoso, la ausencia de ascensos, las condiciones de trabajo inferiores), y, de este modo, existe una mayor variedad de juicios por discriminación posibles. Por lo tanto, una disminución de la discriminación en el empleo puede generar efectivamente un aumento de los juicios por discriminación en el empleo.[14]

XXVIII.5. La discriminación inversa

Se sostiene a menudo que debiera otorgarse a los negros un tratamiento preferente; por ejemplo, que las escuelas de derecho fijaran estándares de admi-

[13] Véase Derek A. Neal y William R. Johnson, "The Role of Premarket Factors in Black-White Wage Differences", 104 *J. Pol. Econ.,* 869 (1996).

[14] Acerca de si el efecto neto de las leyes de discriminación en el empleo ha sido benéfico para los negros, véase la juiciosa discusión de John J. Donohue III y James Heckman, "Continuous *versus* Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks", 29 *J. Econ. Lit.,* 1603 (1991). Estos autores descubren pocos de tales beneficios fuera del Sur.

sión menos exigentes para los negros que para los blancos, aunque los criterios de admisión provean estimaciones no sesgadas del desempeño académico de los negros. Muchas universidades y escuelas de derecho hacen esto. ¿Es tal discriminación inversa algo fundamentalmente diferente de la antigua discriminación *en contra de* los negros? La respuesta a esta interrogante nos obligará a examinar el supuesto empleado hasta ahora de que la discriminación es simplemente un resultado de los gustos e inquirir más a fondo sobre sus causas.[15]

La discriminación racial tiene varias causas posibles. La ignorancia, la maldad y la irracionalidad son factores en muchos casos. La discriminación es a veces anticompetitiva —éste parece haber sido uno de los factores de la reclusión de residentes japoneses en California durante la segunda Guerra Mundial e históricamente ha desempeñado un papel de primera línea en el antisemitismo— y a veces explotadora, como en la esclavitud, en la cual la raza interviene como un factor conveniente para identificar los miembros del grupo rival o explotado. Con todo, otro factor es el de los costos de la información. En la medida en que la raza o algún atributo similarmente difícil de ocultar (el sexo, el acento, etc.) se correlaciona positivamente con la posesión de características indeseables, o se correlaciona negativamente con las características deseadas, entonces, como vimos antes (en XII.10), resulta racional que la gente utilice el atributo como una aproximación a la característica subyacente con la que se correlaciona. Los economistas llaman a esto "discriminación estadística".[16] Al parecer, este fenómeno es común en los mercados de consumidores.[17] Si la experiencia me ha enseñado (quizás incorrectamente)[18] que la mayoría de los originarios de Micenas tienen un fuerte olor a ajo en su aliento, puedo economizar en los costos de información absteniéndome de ingresar a un club que acepte como miembros a tales personas. Aunque podría estar sacrificando asociaciones valiosas con los originarios de Micenas que no tienen ese fuerte olor a ajo, este costo de oportunidad podría ser me-

[15] Véase también antes XXII.10.

[16] Véanse, por ejemplo, David H. Autour y David Scarborough, "Does Job Testing Harm Minority Workers? Evidence from Retail Establishments", 123 *Q. J. Econ*., 219 (2008); Kerwin Kofi Charles y Jonathan Guryan, "Studying Discrimination: Fundamental Challenges and Recent Progress", 3 *Ann. Rev. Econ.*, 479 (2011).

[17] Véanse, por ejemplo, Andrew Hanson y Zachary Hawley, "Do Landlords Discriminate in the Rental Housing Market? Evidence from an Internet Field Experiment in US Cities", 70 *J. Urban Econ.,* 99 (2011); John A. List, "The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field", 88 *Q. J. Econ*., 300 (2004).

[18] Debido a la dificultad de establecer derechos de propiedad en la información, los individuos pueden tener incentivos poco adecuados para investigar aun las características medias de los grupos con los que tratan. Tal vez al forzar a los empleadores a contratar negros podría el gobierno educar a aquéllos sobre lo concerniente a las verdaderas características promedio de los empleados de color, que podrían ser más elevadas de lo que creen quienes dan trabajo. Y, quizás al obligar a los empleadores a dar la misma paga a los negros que a los blancos, el gobierno podría alentar a los negros más productivos a invertir más en su capital humano.

nor al costo de información que implicaría un muestreo más extenso de los originarios de Micenas. Las pruebas de que los solicitantes de empleo con nombres a todas luces propios de los negros reciben menos llamadas de respuesta que los solicitantes con nombres "normales"[19] guarda correspondencia tanto con la discriminación basada en los gustos personales como con la estadística. Adviértase también que, debido a la existencia de una correlación positiva entre, por un lado, la raza y, por otro, las características que tienen un peso en las decisiones de los empleadores, prestamistas, etc., se torna imposible inferir la presencia de pautas discriminatorias en los resultados del mercado, tales como el hecho de que, en promedio, los negros confrontan costos de préstamo en bienes raíces más elevados que los blancos.[20]

Que alguna discriminación racial resulte eficiente no significa que sea o deba ser legal. Por razones utilitarias podría ser injusta a causa de sus efectos distributivos, aun cuando sea eficiente. Supongamos, por ejemplo, que todos los secuestradores de aviones fuesen miembros de un grupo étnico particular, pero que sólo un porcentaje pequeño de los miembros del grupo fueran secuestradores de aviones. Podría ser una estrategia policiaca racional utilizar un "perfil étnico" de tal modo que sólo se registrara a los miembros del grupo en cuestión antes de que se les permitiera abordar un avión. Pero entonces *todo* el costo de los registros de personas inocentes en los aeropuertos sería soportado por los miembros de un grupo étnico.[21]

Por más que se ponderen fuertemente las consideraciones distributivas, no cabe exceptuar cierto uso del perfil étnico en virtud de que resulta demostradamente eficiente. Si los testigos reportan un robo cometido por un varón joven de raza negra, sería absurdo que la policía buscara sospechosos entre miembros de otros grupos demográficos, aun si el resultado fuese que se sometiera a un interrogatorio a muchos negros inocentes y a ningún blanco. Cuando el diferencial de la raza en la probabilidad de la culpa es mucho menor, el perfil étnico deviene problemático incluso por razones de pura eficiencia. Fijémonos en una política basada en el cateo desproporcionadamente constante de vehículos conducidos por hispanos en virtud de que es desproporcionado el número de hispanos envueltos en el tráfico de drogas ilegales. La política en cuestión disuadirá a algunos hispanos de ingresar al comercio de estupefacientes o de seguir en él. Pero el asunto es que obligar a los miembros de una clase de proveedores a dejar el negocio deja el espacio abierto para que otros entren en él. Dado un presupuesto fijo para aplicar la ley en este terreno,

[19] Marianne Bertrand y Sendhil Mullainathan, "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination", 94 *Am. Econ. Rev*., 991, 992 (2004).

[20] Harold A. Black, Thomas P. Boehm y Ramon P. DeGennaro, "Is There Discrimination in Mortgage Pricing? The Case of Overages", 27 *J. Banking & Fin.,* 1139 (2003).

[21] Véase Amy Farmer y Dek Terrell, "Crime *versus* Justice: Is There a Trade-Off?", 44 *J. Law & Econ.,* 345 (2001).

el incremento en la aprehensión de transportistas de la droga de origen hispano quedaría contrarrestado por una reducción en el riesgo que enfrentan las personas de origen no hispano de ser aprehendidas por llevar a cuestas narcóticos, y, de esta manera, los que no son hispanos habrán de acudir en tropel a remplazar a los transportistas hispanos. La composición étnica de la fuerza de trabajo ilegal sufriría una alteración, pero la tasa de delitos sería afectada exclusivamente en la medida en que los hispanos son más eficaces para llevar drogas de un lugar a otro merced a su dominio de la lengua y a otros vínculos que tienen con los principales países productores de narcóticos. A la inversa, entre más baja sea la elasticidad cruzada de la oferta entre el grupo perfilado y el no perfilado, más probabilidades habrá de que el perfil étnico funcione con eficiencia porque habrá menos sustitución de miembros del grupo no perfilado una vez que los recursos de aplicación de la ley dejen de estar destinados a ellos y se reasignen a los del perfilado.

Cuando el perfil étnico produce la misma tasa de éxito a través de los grupos raciales —y existen evidencias de que el perfil de esta índole usado por la policía hoy en día realmente alcanza este logro—,[22] las objeciones distributivas a él se desvanecen, y el acto de prohibir semejante perfil bien podría juzgarse una clase de discriminación a la inversa. Y, de manera similar, si los delincuentes que pertenecen al grupo bajo perfil son menos capaces de experimentar la disuasión que los demás grupos, hay bases para defender un argumento neutral a la raza a favor de aprehender a una mayor porción de ellos, tanto con el propósito de aumentar la disuasión como con el de dejar maniatados a más sujetos imposibles de ser disuadidos. Y si los miembros del grupo perfilado son sencillamente más fáciles de aprehender, entonces aprehenderlos a una tasa más elevada conducirá a reducir los delitos a un costo menor.[23]

Lo mismo en el área del perfil étnico que en el de la discriminación estadística, la raza se utiliza como una aproximación a las características personales subyacentes, y cabe sostener que, a menos que la Decimocuarta Enmienda y otras leyes antidiscriminatorias se interpreten de manera tal que se prohíba el uso en cuestión, la protección legal de las minorías raciales será magra. La implicación, empero, es que la discriminación inversa es anticonstitucional, debido a que ella también está basada en el uso de la raza a manera de una aproximación a las características personales subyacentes. No puede estar fundada racionalmente en un deseo de proporcionar reparaciones a los

[22] Véanse Nicola Persico, "Racial Profiling, Fairness, and Effectiveness of Policing", 92 *Am. Econ. Rev.*, 1472 (2002); Nicola Persico y David A. Castleman, "Punishment and Crime: Detecting Bias: Using Statistical Evidence to Establish Intentional Discrimination in Racial Profiling Cases", 2005 *U. Chi. Legal Forum,* 217; Nicola Persico y Petra E. Todd, "Passenger Profiling, Imperfect Screening, and Airport Security", 95 *Am. Econ. Rev.*, 127 (2005); Anderson *vs.* Cornejo, 355 F.3d, 1021, 1024 (7th Cir., 2004).

[23] Véase Anderson *vs.* Cornejo, nota 22, en 1025.

negros respecto a injusticias provocadas en el pasado, como la esclavitud, el linchamiento y las leyes de Jim Crow. No existe forma alguna de compensar a los difuntos, ni tampoco la hay de determinar la posición económica que alcanzarían los negros en la actualidad si no hubiera sido por esos agravios.[24]

La discriminación inversa trae consigo una redistribución caprichosa del ingreso *entre* los negros. En la medida en que los empleadores encuentren costoso evaluar las cualidades de un trabajador en lo individual, los trabajadores negros que, ubicados en una línea de trabajo en particular, son igual de buenos que los blancos en la misma línea encontrarán difícil desligarse de la discriminación positiva y, por consiguiente, se dará por sentado que poseen las cualidades promedio de los trabajadores negros en esa línea de trabajo… un promedio que ha sido arrastrado hacia abajo por las contrataciones derivadas de la discriminación afirmativa. Para formularlo de acuerdo con la teoría de juegos, habrá un equilibrio de agrupación en el que los trabajadores negros ubicados por debajo del promedio saldrán beneficiados a expensas de quienes están por encima del promedio.

Los costos de información ayudan a explicar también la razón de que sea tan impopular la discriminación inversa. Cada vez que un varón blanco no logra ser admitido en una universidad de la que se sabe que practica la discriminación inversa, o que no es contratado por un empleador de quien se sabe que la practica, hay alguna probabilidad de que haya perdido esta oportunidad valiosa por causa de la discriminación inversa. Pero la probabilidad será a menudo mucho menor a uno. Supongamos que cuatro personas solicitan un empleo. Tres son blancas y una es negra, y la negra es contratada. Aunque todos los blancos sean mejores que el negro, quien fue contratado sólo gracias a su raza, sabemos que dos de los tres blancos no habrían obtenido el empleo aunque el empleador siguiera una política de indiferencia ante el color. No obstante, es posible que estos dos no sepan quiénes son; por lo tanto, los tres blancos creerán que han sido perjudicados por la discriminación inversa. Como si dijéramos, cada uno de ellos es una víctima probabilística de la práctica.

Lecturas recomendadas

Becker, Gary S., *The Economics of Discrimination* (2ª ed., 1971).

Discrimination in Labor Markets (Orley Ashenfelter y Albert Rees [comps.], 1973).

Donohue, John J. III, "Is Title VII Efficient?", 134 *U. Pa. L. Rev.,* 1411 (1986).

——— y Peter Siegelman, "The Changing Nature of Employment Discrimination Litigation", 43 *Stan. L. Rev.,* 983 (1991).

[24] Véase In re African-American Slave Descendants Lit., 471 F.3d, 754 (7th Circ., 2006).

Farmer, Amy y Dek Terrell, "Crime *versus* Justice: Is There a Tradeoff?", 44 *J. Law & Econ*., 345 (2001).
Persico, Nicola, "Racial Profiling, Fairness, and Effectiveness of Policing", 92 *Am. Econ. Rev*., 1472 (2002).
Phelps, Edmund S., "The Statistical Theory of Racism and Sexism", 62 *Am. Econ. Rev.,* 659 (1972).
Philipson, Tomas, "Desegregation and Social Monopoly Pricing", 4 *Rationality and Society,* 189 (1992).
Posner, Richard A., "The Efficiency and the Efficacy of Title VII", 136 *U. Pa. L. Rev.,* 513 (1987).
Sowell, Thomas, *Civil Rights: Rhetoric or Reality?* (1984).

Problemas

1. Este capítulo ha sugerido un principio neutral para prohibir la discriminación. ¿Es un principio económico? ¿Podemos afirmar que la discriminación es ineficiente? En términos económicos, ¿son los costos de las asociaciones interraciales, dado el prejuicio, diferentes en cualquier grado del daño causado a los cultivos por la interacción del ferrocarril y la agricultura?

2. Supongamos que varios negros compraron casas con contratos de tierras y luego los incumplieron. Sostienen que no debieran ser declarados culpables de no pagar su deuda porque se vieron obligados a pagar precios mayores a los de los compradores blancos por propiedades similares, como resultado de la discriminación contra los negros. Los desarrolladores replican que los negros debieran estar agradecidos de que ellos estuvieran dispuestos a venderles una propiedad tan valiosa. ¿Cómo puede el análisis económico aclarar los problemas de tal litigio? ¿Aumentaría o disminuiría el bienestar de los negros en conjunto si los desarrolladores perdieran?

3. ¿Puede sostenerse que los convenios racialmente restrictivos podrían aumentar la eficiencia?

4. Supongamos que una escuela de derecho que descubriera que sus graduados negros tenían ingresos profesionales de toda la vida menores a los de los blancos, a causa de la discriminación racial, decidiera imponer en consecuencia requisitos de admisión más exigentes a los negros que a los blancos. ¿Podría defenderse esta política porque aumenta la eficiencia?

5. Los varones negros tienen una esperanza de vida más corta que los varones blancos. Discuta los efectos de asignación y distribución de las reglas que prohíben que las compañías de seguros de vida varíen las tasas de sus primas sobre la base de la raza del asegurado.

6. Compare dos formas de la discriminación inversa: en una de ellas, el empleador fija una cuota para los empleados negros y contrata sólo negros hasta que se llene la cuota; en la otra, el empleador contrata sin discrimina-

ción pero otorga a sus empleados negros una antigüedad mayor que a sus empleados blancos, de modo que cuando las condiciones económicas lo requieran despide a menos negros que blancos. Considere si hay alguna diferencia en el hecho de que el empleador sea público o privado, de que exista o no un sindicato, y de que la política de otorgar una antigüedad mayor a los negros se adopte antes o después de la contratación de los trabajadores blancos afectados por ella. ¿Cuál combinación de atributos produce la discriminación más ineficiente? ¿Cuál es la menos ineficiente?

7. Los empleadores que desean discriminar a los negros, pero la ley se los prohíbe, podrían buscar una característica aproximada que poseen más negros que blancos y usarla como la base de las decisiones de personal. La aproximación podría ser algún nivel de logro educativo. Si usted es un empleador, ¿se preocuparía más porque se le prohíba el uso de la aproximación o porque se le exija tener una fracción específica de empleados negros? Véase Shelly J. Lundberg, "The Enforcement of Equal Opportunity Laws under Imperfect Information: Affirmative Action and Alternatives", 106 *Q. J. Econ*., 309 (1991).

8. En un estudio empírico —Ian Ayres, "Fair Driving: Gender and Race Discrimination in Retail Car Negotiations", 104 *Harv. L. Rev.,* 817 (1991)— se encuentra que las mujeres y los negros pagan en promedio precios más elevados por los mismos automóviles que los varones blancos, y se atribuye este fenómeno a un modelo de negocios de "búsquese a los cándidos que se lo tragan todo" adoptado por los vendedores de autos: la mayor parte de las ganancias de estos últimos derivan de la minoría de ventas con un margen de beneficio o sobreprecio muy elevado con respecto al costo del comerciante (esto es, se trataba de ventas cercanas al precio inicial de catálogo del automóvil). ¿Por qué debería un número desproporcionado de incautos corresponder a negros o a mujeres, y cuál es la lógica económica de un modelo de negocios de "búsquese a los cándidos que se lo tragan todo"? (Formulado de otra manera, ¿un mercado que emplea semejante modelo mercantil constituye un equilibrio competitivo?)

XXIX. LA PROTECCIÓN DE LOS MERCADOS LIBRES EN LAS IDEAS Y EN LA RELIGIÓN

XXIX.1. La base económica de la libertad de expresión

Las ideas son un bien útil producido en cantidades enormes en un mercado muy competitivo. El mercado de las ideas del que escribiera Holmes es un hecho, no sólo un modo de hablar.[1] Desde el punto de vista práctico, es este mercado, antes que alguna realidad trascendental, lo que determina la "verdad" de las ideas. Decimos que una idea (por ejemplo, que la Tierra gira alrededor del Sol) es cierta no porque sea *realmente* cierta —¿quién lo sabe?—, sino porque todos los consumidores expertos, o la mayoría de ellos, la han aceptado ("comprado"). Esta concepción pragmática de la verdad[2] mina los esfuerzos de supresión de las ideas o de prohibición de su expresión o difusión. Nadie tiene un conducto directo hacia la realidad trascendente. Las ideas que poseemos se forjan en un proceso competitivo que se distorsiona si se excluye por la fuerza a los competidores potenciales: las ideas impopulares o repulsivas.

Pero esto no explica por qué se ha considerado necesaria la protección constitucional para este mercado particular y no para otros. Dos explicaciones posibles concuerdan con el pensamiento económico. La primera es que la intervención pública en el mercado de las ideas genera un peligro de subversión del proceso democrático, lo que conduce al que es el más peligroso de los monopolios: el monopolio del poder gubernamental (véase antes XXV.2). Y la historia enseña que la democracia, con todas sus imperfecciones, es enormemente superior a los regímenes autoritarios (con sus excepciones: muchas naciones aún parecen carecer de las condiciones económicas y culturales que deben anteceder a una democracia.). La segunda y más amplia forma de explicar la protección constitucional de un mercado libre (bueno, relativamente libre) de ideas y de opinión (la primera se limita a la expresión política) destaca la fragilidad de los mercados de la información. Por las razones exploradas antes en el capítulo XI, no se pueden crear derechos de propiedad en las ideas puras, consideradas como distintas de las invenciones prácticas. Por lo tanto, es probable que se produzcan en una cantidad deficiente. El problema es particularmente grave debido a que las ideas populares poseen una ventaja en la competencia contra ideas valiosas pero impopulares. Entonces cualesquiera

[1] Abrams *vs.* United States, 250 U.S., 616, 630 (1919) (opinión disidente).

[2] Véanse, por ejemplo, Richard Rorty, *Philosophy and the Mirror Nature* (1979); Thomas S. Kuhn, *The Structure of Scientific Revolutions* (4ª ed., 2012). [Hay traducción del FCE en Breviarios.]

costos que el gobierno imponga a las ideas impopulares podrían causar su abandono en masa. En efecto, la conjunción de lo "valioso" y lo "impopular" sugiere la existencia de una clase de ideas cuyos beneficios son casi enteramente externos. Por lo tanto, hay una razón económica para preocuparse por el "congelamiento" del ejercicio de la libertad de expresión.

El argumento de los beneficios externos para otorgar al mercado de las ideas una protección legal especial no se limita a la expresión política, pero esa expresión puede ser su mejor parangón. Recordemos que la votación misma es una fuente de beneficios externos, porque un solo voto decidirá raras veces una elección, de modo que su valor esperado para el votante individual es pequeño aun en relación con los escasos costos de tiempo de la votación. Dado que la votación tiene escaso valor privado, no debiéramos esperar que la gente invierta fuertemente para informarse acerca de los candidatos y los temas. Esto significa que la demanda privada de ideas y opiniones políticas tenderá a ser débil, lo que es un argumento contundente a favor de la minimización de los costos legales de la producción de tales ideas y opiniones.

¿Pueden las restricciones al financiamiento de las campañas reconciliarse con la subinversión, implicada por el análisis expuesto arriba de los beneficios externos, en la creación y difusión de las ideas y opiniones políticas? Prohibirle a una persona gastarse más de, digamos, un millón al año para publicitar sus ideas políticas sería en todo punto incongruente con el análisis, y, con base en ello, la Corte Suprema ha dictaminado que a un candidato a un puesto político no puede prohibírsele gastar todo el dinero salido de su bolsillo que quiera en su propia campaña. Le ha permitido al gobierno fijar límites a cuánto puede uno aportar a la campaña de otra persona; esos límites hacen falta para atenuar los efectos distorsionadores que tienen sobre las políticas públicas los grupos de interés con sus consignas. Sin embargo, en lo que parece contradecir su medida de poner restricciones a las contribuciones hechas a una *campaña,* recientemente ha invalidado una provisión de la Ley McCain-Feingold (véase antes XX.7) que prohibía a las corporaciones (incluidas las no lucrativas), así como a los sindicatos, patrocinar anuncios a favor o en contra de candidatos en los 30 días previos a unas elecciones primarias y en los 60 días previos a unas elecciones generales, a condición de que los patrocinadores fuesen independientes de la organización de las campañas.[3] El peligro reside en que tales gastos, concentrados en el periodo crítico de una campaña, podían, especialmente en un distrito electoral de pequeñas dimensiones, representar un peso abrumador sobre la publicidad del candidato. Al percatarse de ello, los candidatos tendrán un incentivo extra para inclinarse servilmente a las plantas de los integrantes de los grupos de interés más influyentes. Y reparemos en que el costo que tiene para estos grupos la obtención de una influencia adicional puede ser bajo o incluso nulo; pues tan sólo la amenaza de

[3] Citizens United *vs.* Federal Election Commission, 130 S. Ct., 876 (2010).

una intervención financiera masiva en los días de cierre de campaña puede ser suficiente para disuadir a un candidato de prestar su apoyo a una política que sería un anatema para los grupos de interés más poderosos.

La cuestión es si el costo del incremento del ascendiente que posee el grupo de interés será compensado por el incremento en la información derivado de un aumento en la publicidad política. La decisión de la Corte Suprema, emitida en 2010, puede haber contribuido al aumento de los gastos en que incurrieron tales grupos al pasar de 157 millones de dólares en las elecciones federales de 2008 a más de mil millones en las elecciones de 2012.[4]

XXIX.2. El alcance de la actividad protegida: incitación, amenazas, difamación, obscenidad y quema de la bandera

Podría parecer que la teoría pragmática de la verdad, aunada a la observación recién apuntada acerca de los beneficios externos, apunta hacia una prohibición absoluta de la regulación pública de la expresión. No es así. La expresión tiene muchas funciones además de la comunicación de ideas. Y las ideas, ciertas o falsas, pueden causar un daño que justifique la intervención pública. Si digo "Voy a dar un paseo" o "Voy a robar un banco" o "Estoy organizando una insurrección armada", no estoy apelando al mercado de las ideas, sino sólo declarando una intención, y mi declaración puede ser una prueba de un intento de cometer un delito que quizá no tenga nada que ver con las ideas, un delito tal como el robo. El castigo del intento no obstruye el mercado de ideas en absoluto. Las declaraciones de intención no tratan de competir con otras ideas, como una declaración de que la tierra es plana trata de competir con otras opiniones, y no hay ningún riesgo de una producción insuficiente porque no hay ninguna inversión en la producción de la idea que se encuentra detrás de la declaración de intenciones.

Supongamos ahora que digo: "La insurrección armada mañana sería una buena idea" o (si soy un productor de adminículos) "La industria estaría mejor si el precio de los adminículos fuese 10% mayor" o "Intento votar por x". Estas declaraciones expresan ideas genuinas, porque hacen un esfuerzo por desplazar las ideas rivales en el mercado de las ideas. El problema es que las dos primeras podrían ser también invitaciones a cometer actos ilegales (traición y fijación de precios, respectivamente). Como invitaciones, parecerían punibles según el mismo principio que vuelve punibles los atentados y las conspiraciones. Pero el castigo tendrá el efecto colateral de suprimir una idea.

Una fórmula económica para el tratamiento de estos casos mixtos de idea e incitación fue propuesta por el juez Learned Hand (por supuesto) en United

[4] Samuel Issacharoff y Jeremy Peterman, "Special Interests after Citizens United: Access, Replacement, and Interest Group Response to Legal Change", 21-22 (N.Y.U. Law & Econ. Working Paper 328, 2013), http://lsr.nellco.org/nyu_lewp/328.

States *vs.* Dennis.[5] Los tribunales —escribió Hand— deben en cada caso "preguntarse si la gravedad del 'mal' [es decir, si triunfa la instigación], descontada por su improbabilidad, justifica la invasión de la libertad de expresión que es necesaria para evitar el peligro".[6] Esto equivale a la Fórmula de Hand para la negligencia ($B < PL$) si se define B como el costo de la disminución del acervo de ideas como resultado de la acción gubernamental, P como la probabilidad de que se cometan los delitos aconsejados por el declarante, y L como el costo social de que en verdad se cometan. Si B es menor a PL, será eficiente que el gobierno tome medidas contra el declarante.

La fórmula de Dennis es una expansión de la prueba de Holmes, más famosa, de un "peligro claro y actual".[7] La diferencia se genera cuando modificamos la Fórmula de Hand para tomar en cuenta el hecho de que el peligro podría encontrarse en el futuro, como ocurrió en Dennis, un enjuiciamiento de los líderes del Partido Comunista de los Estados Unidos por complotar para derrocar al gobierno… a la larga. Si i es la tasa de descuento de los daños futuros al presente y n es el número de años que deben transcurrir hasta que ocurra el daño, entonces $B < PL$ se vuelve $B < P \times L / (1 + i)^n$. Entre mayor sea la tasa de descuento y más remoto sea el daño, menor será el miembro derecho de la fórmula y, por ende, más débil será la justificación de la supresión. Sin embargo, si L es suficientemente grande, todavía podría haber una justificación respetable para la supresión, aunque n sea sustancial. Pero no de acuerdo con la fórmula de Holmes, que descarta todos los casos donde n sea positivo. Eso parece arbitrario. Y adviértase el hecho curioso de que, entre más "preocupados por el futuro" seamos como sociedad, más estaremos dispuestos a suprimir las ideas peligrosas aunque, como en el caso de las predicaciones de los comunistas estadunidenses, el peligro se encuentre, si acaso, en el futuro remoto. Pues entre más volcados hacia el futuro estemos, más baja será i, la tasa social de descuento, y, en consecuencia, más grande será el miembro derecho de la fórmula de Dennis expandida. (Colocados en el límite, si i es cero, el descuento no tiene efecto alguno sobre L, puesto que la fórmula se reduce a $B < P \times L$.)

La aplicación de la fórmula de Dennis depende justamente de los pasos que esté dispuesto a dar el gobierno. Si se propone castigar penalmente al declarante, B será sustancial y requerirá en consecuencia una PL sustancial para contrarrestarlo. Pero si el gobierno sólo se propone monitorear la actividad del declarante, de modo que podrá actuar cuando se vuelva inminente el peligro de una violación delictiva, B será menor porque la desistencia de la libre expresión será menor y, por ende, una PL menor que en el primer caso bastará para contrarrestar a B y justificar la acción gubernamental.

[5] 183 F.2d 201 (2d Cir., 1950), *aff'd,* 341 U.S., 494 (1951).
[6] *Ibid.,* en 212. Quería decir "descontada por su probabilidad".
[7] Schenck *vs.* United States, 249 U.S., 47, 51 (1919).

La fórmula, por imposible de cuantificar que sea, resulta útil para explicar la razón de que, por ejemplo, es más probable que se tolere la argumentación a favor de males muy grandes —el genocidio, la revolución o lo que sea— a que se defienda el llamado a un linchamiento, que es un mal menor, o aun la comisión del "mal" trivial creado por un ruidoso camión de sonido. Si las circunstancias vuelven remota la probabilidad de que triunfe la propaganda genocida, el costo descontado de la declaración podría ser menor al de una amenaza de linchamiento. En el caso del camión de sonido, son pequeños los daños causados por su estruendo (*L* en la fórmula), pero también lo es el costo de los beneficios sacrificados, ya que el declarante puede propagar su mensaje por medios menos ofensivos. El caso del camión de sonido, como otros casos que implican restricciones al tiempo, lugar y modo, antes que a la sustancia del discurso, es analíticamente similar a nuestro caso del gobierno que sólo investiga al declarante en lugar de castigarlo. Adviértase que el camión de sonido y la incitación al delito imponen costos externos, lo que constituye una justificación tradicional de la regulación.

La fórmula de Dennis podría parecer paternalista. Supongamos que un grupo está tratando de persuadir a la gente de que una revolución violenta mejoraría su situación, y las circunstancias vuelven la probabilidad del éxito suficientemente elevada para desatar la prueba, aunque no se esté aconsejando ninguna acción revolucionaria inmediata. Dado que hay tiempo para que grupos rivales persuadan a la gente de que una revolución no mejoraría su situación, ¿para qué interferir con el mercado de las ideas? Una respuesta es que, dado el intervalo existente para la persuasión contraria, *P* es realmente muy pequeña, de modo que la fórmula no justificaría la represión. La justificación de la represión es más fuerte cuando, como ocurre en el caso habitual de la incitación, el intervalo entre el discurso y la acción es demasiado breve para permitir que se presenten concepciones rivales. En tal caso, castigar el discurso es como castigar los monopolios situacionales (como el de los marineros en el caso Alaska Packers, antes en IV.2): hay una clase similar de falla del mercado. Tampoco en el caso del camión de sonido puede confiarse en que el mercado de las ideas proteja a las víctimas del daño, porque los costos son externos al mercado de las ideas. Un segundo camión de sonido empeoraría las cosas en lugar de mejorarlas.

Hay un argumento económico similar para suprimir la defensa de la revolución violenta aun en el futuro distante. Aunque tal defensa podría contener ideas (por ejemplo, que los capitalistas obtienen ganancias mayores de las que debieran, o que se está ampliando la brecha entre ricos y pobres), es también una invitación a realizar actividades que son delictivas. La invitación podría ser atractiva incluso si el mercado de las ideas demostrara convincentemente la falsedad de las ideas del defensor. Supongamos que el declarante insta a los pobres a levantarse y apoderarse del dinero de los ricos porque éstos los están explotando. Aun si la propaganda contraria demostrara con-

vincentemente que los ricos no están explotando a los pobres, sigue siendo cierto que los pobres podrían optar por levantarse y despojar a los ricos, como los incita a hacerlo el propagandista. Podría pensarse que tal discurso no es otra cosa que una invitación a "robar un banco", sólo agrandada.

Hasta ahora, la discusión ha supuesto que existe una línea clara entre una declaración de intenciones y una idea. Pero consideremos la clase de declaraciones de intención conocidas como amenazas. "Digo que te volaré la cabeza si no me entregas tu cartera." La amenaza comunica una información genuina acerca de mis intenciones… pero sólo en cierto sentido; porque está en la naturaleza de la mayoría de las amenazas que el que la expresa no desea cumplirlas si no es para mantener la credibilidad. Por lo tanto, si las amenazas se suprimen eficazmente, también se suprimirán (de ordinario) los males sustantivos con los que se amenaza. De igual modo, la inversión hecha en la formulación de la amenaza no tiene ningún producto social; por lo tanto, deseamos desalentarla, no protegerla. Pero veamos ahora el caso en el que amenazo con hacer algo completamente legal, como "entregarte a la policía si no me pagas el silencio". ¿Por qué debiera una declaración de intención volver punible un acto legal? No obstante, de nuevo la clave es que la intención es condicional. "No deseo entregarte, quiero tu dinero." Por lo tanto, si el chantaje es ilegal (véase antes XXIV.2), las peticiones que se hacen cuando se comete el chantaje debieran ser ilegales también. Es poco importante que la petición asuma la forma de una zanahoria ("Te ofrezco mi silencio por dinero") o un garrote ("Si no me pagas, te entregaré").

Hablando de chantaje, debemos señalar que las declaraciones acerca de los individuos son una parte importante del mercado de las ideas y en principio pueden ser validadas o refutadas en la misma forma que otras ideas: por la competencia. En los términos de la Fórmula de Hand, *B* es el costo de evitar la falsedad (un costo que incluye cualquier pérdida de beneficios sociales a causa de la actividad de investigación), *L* el daño provocado a la persona a quien se difamó y *P* la probabilidad de que la declaración sea de hecho falsa. En virtud de que la falsedad de la difamación es fácilmente demostrable, puede encontrarse en una determinación legal de la verdad un buen sustituto de una determinación realizada por el mercado. Y la competencia puede no ser un remedio efectivo: ¿cómo ha de competir un individuo oscuro, venido de no se sabe dónde, con Fox News si la compañía le entabla un juicio por libelo?

Así pues, los productores y vendedores de ideas, así como los de otros bienes, debieran ser responsables por perjuicios causados a una reputación. Empero, la Corte Suprema, pronunciándose en nombre de la Primera Enmienda, ha impuesto estrictas limitaciones a los remedios legales por difamación. En particular, una figura pública (por lo regular, aunque no siempre, un político o un representante oficial, así como puede serlo un deportista célebre o alguien bien conocido del mundo de los espectáculos) puede no recuperar los daños a menos que le sea posible probar que el acusado sabía que la difamación era

infundada o, por proceder precipitadamente, fracasó en el intento de determinar si lo era. Las personas que no son figuras públicas poseen derechos a entablar juicios que son algo más amplios. La distinción tiene sentido económico. Una figura pública cuenta con mejores oportunidades de utilizar los medios para llevarle su respuesta al difamador que una figura no pública, y una contestación es superior a un juicio por difamación desde la perspectiva de la Primera Enmienda debido a que preserva el mercado de ideas sin regulación de por medio. Asimismo, la publicación de críticas de figuras públicas confiere beneficios informativos que no van a parar totalmente a manos del editor.

El efecto que provoca el requisito de probar que se incurrió en desconsideración intencional o por descuido consiste en excusar al editor de la responsabilidad por mera negligencia, y hemos visto otros casos donde fue esto lo que ocurrió porque el demandado había conferido un beneficio externo mediante el acto negligente (véase, por ejemplo, antes VI.4). Si un reportero obtiene un pitazo, su periódico captará en mayores ingresos por ventas sólo una parte del valor que el público asigna a la noticia, porque ésta se publicará en todos los periódicos rivales al poco tiempo. Por lo tanto, si el reportero y el periódico que lo emplea afrontan la perspectiva de grandes daños, podrían resistirse a publicar la historia aunque los beneficios sociales totales (pero no sus beneficios privados), medidos por la disposición de todos los lectores de periódicos a pagar por leer la noticia, pudieran superar tales daños. Un procedimiento para alentar al periódico a publicar la noticia consiste en disminuir los costos de su publicación; y esto se hace volviendo innecesario que el periódico realice una investigación de la verdad de la noticia tan completa como tendría que hacerlo si fuese estrictamente responsable o responsable por negligencia al publicar una difamación falsa.

Una objeción a ese enfoque es que obliga a las víctimas de la difamación a subsidiar la producción de ideas. Esto podría no ser tan malo si fuese fácil asegurarse contra las consecuencias de ser difamado, porque entonces la incidencia del subsidio se dispersaría entre todo el grupo asegurado. No es fácil. La forma actual del subsidio tiene el efecto curioso de desalentar a la gente de convertirse en una figura pública. Un enfoque alterno consistiría en subsidiar directamente a los medios pero permitiendo que las víctimas de la difamación conservaran completos sus derechos comunes… aunque sería probablemente imposible determinar cuán abundante habría de ser el subsidio que se le asignara. Otro enfoque sería el de reconocer más extensos derechos de propiedad en las noticia, como el caso de la decisión de Associated Press,[8] donde se sostuvo que un servicio de noticias podría demandar por la publicación no autorizada de sus despachos por parte de un servicio rival.

[8] International News Service *vs.* Associated Press, 248 U.S. 215 (1918). La decisión, tomada según la ley del derecho común federal, ya no es oficial, pero hay una serie de estados que reconocen su principio, a saber, que la apropiación indebida de la inversión realizada por otra persona en información puede ser perjudicial.

En los casos del discurso peligroso y de la difamación hay una externalidad dañina que provee un argumento económico convencional para la regulación: el orador incita a sus oyentes a herir a otros, o el periódico enciende a sus lectores acusando a alguien de un comportamiento escandaloso. También la pornografía crea una externalidad, al menos cuando se muestra públicamente a una audiencia renuente, como ocurre en las marquesinas de los teatros de Londres. La supresión del despliegue público de la pornografía elimina la molestia a un costo mínimo, medido por la disminución de la eficacia de la publicación de anuncios pornográficos por medios no pornográficos. Éste es un buen ejemplo de las restricciones de tiempo, lugar y modo para la libre expresión, las que tratan con mayor benignidad que la supresión directa. Son similares las regulaciones de zonificación que confinan a las tiendas encargadas de vender materiales o entretenimiento pornográficos a distritos no residenciales especificados. Si el gobierno regula los lugares en donde se puede apreciar el arte sexualmente explícito, hace que disminuya la audiencia para el arte y, a la vez, el incentivo para crearlo en primer lugar. Pero lo lleva a cabo sólo en un grado menor (dependiendo de la naturaleza precisa de la regulación), en tanto que si le aplica un castigo propio de un delincuente a quienquiera que cree dicho arte reduce en gran escala el incentivo para crearlo.

La mayoría de quienes se oponen a la pornografía no están satisfechos con la regulación: desean la prohibición total. Consideremos los argumentos esgrimidos por Catharine MacKinnon contra la pornografía:[9] *1)* La pornografía incita a sus consumidores a cometer violaciones sexuales. *2)* Los modelos y las actrices (quizá también los actores) son abusados y explotados con frecuencia por los productores de libros y películas pornográficos. *3)* A pesar de la Primera Enmienda, los tribunales permiten que el despliegue de la pornografía —y aun la mera insinuación sexual verbal— en el lugar de trabajo sea presentado como prueba del acoso sexual. *4)* La pornografía no confiere ninguno de los beneficios sociales asociados a la libertad de expresión; trata de encender sexualmente a sus lectores antes que de informarlos.

La economía apoya la última observación, pero arroja dudas sobre las primeras tres.

1. La pornografía, como vimos antes en V.7, es un complemento (énfasis de MacKinnon) y un sustituto del sexo coercitivo (esto último porque la pornografía es un complemento de la masturbación, que es un sustituto del coito), y no está claro cuál efecto domina.
2. La explotación de las modelos y actrices pornográficas es un artefacto de la ilegalidad de la pornografía. Sabemos por la Prohibición, las leyes contra la prostitución, la campaña contra las drogas y el empleo de inmigrantes ilegales que, cuando se coloca una actividad económica

[9] Véase Catharine A. MacKinnon, *Only Words* (1993).

fuera de la protección del derecho, los participantes recurrirán a las amenazas y la violencia en lugar de los remedios contractuales y de otra clase que se les niegan. El proxeneta que recurre a la fuerza o a la amenaza para ejercer el dominio sobre sus prostitutas es un artefacto de la ilegalidad de la prostitución, y la explotación de las actrices y los modelos pornográficos por sus empleadores es paralela a la explotación de los trabajadores inmigrantes ilegales por sus empleadores. Estas mujeres, por lo menos, estarían en mejor situación si fuera legal toda la pornografía.

3. El tratamiento asimétrico de la pornografía y la insinuación sexual dentro y fuera del lugar de trabajo refleja una diferencia de los costos. Los costos son mayores en el lugar de trabajo porque allí las mujeres son una audiencia involuntaria de la plática o los despliegues que las ofenden. En el segundo caso la pornografía se dirige a un hombre, el lector o espectador de la pornografía, y el objetivo es agradar, no insultar ni intimidar. En virtud de que una mujer no es el espectador o lector buscado y raras veces es el efectivo, sólo puede ser dañada si un consumidor varón voluntario de la pornografía es incitado por esta última a maltratarla. El efecto es indirecto; no está claro si es sustancial (punto 1).
4. La pornografía confiere beneficios a sus consumidores —de otro modo no la comprarían—, pero no confiere beneficios *externos* considerables. Es decir, el pornógrafo, como el dueño de un burdel, debiera ser capaz de captar una parte tan grande del valor de mercado de su producto como el productor en un mercado normal.

La producción y consumo de la pornografía infantil involucra un problema diferente. La producción (dando por sentado que se trata de imágenes fotográficas o de video de niños auténticos en una relación sexual) de esa clase de pornografía causa daños psicológicos a los infantes que fungen de modelos, y el consumo induce a que se realice semejante producción creando un mercado para ella. Tal es la justificación económica para ejercer un castigo penal contra ese consumo y contra la producción, como se acostumbra llevar a cabo. La Corte Suprema ha sostenido que la quema de la bandera estadunidense para expresar una opinión política es una expresión constitucionalmente protegida siempre que el incendiario sea el dueño de la bandera que está quemando. La quema de la bandera es intensamente ofensiva para los estadunidenses patriotas. Puede sostenerse que *PL* es elevada porque *L* no es trivial y *P* es 1 —el daño generado por la quema de la bandera, el de ofender a los patriotas, es una certeza—, y que *B* es baja o aun negativa porque el incendiario de la bandera podría comunicar su mensaje con la misma eficacia, o mayor aún, mediante palabras. Pero este argumento es débil. La eficacia comunicativa de las palabras no es siempre mayor que la de los dibujos u otros sím-

bolos no verbales. El incendiario de la bandera cree claramente que el hecho de quemar una bandera transmite un mensaje más poderoso que si se habla acerca de la inequidad de las políticas estadunidenses. Además, cualquier mensaje antipatriótico igualmente eficaz sería igualmente ofensivo para los patriotas. Por lo tanto, lo que está implícito en una propuesta de prohibir la quema de la bandera es un juicio de que la expresión debe suprimirse cuando causa más dolor que placer, un juicio que si se aplicara congruentemente terminaría con la libertad de expresión tal como la conocemos. Con todo, lo que resulta extraño aquí desde la perspectiva estrictamente legal es que el acto de quemar la bandera no constituye una expresión o discurso en un sentido literal y no lo habrían concebido como constitucionalmente privilegiado quienes formularon o ratificaron la Primera Enmienda.

Incluso las ideas que no tienen ningún valor en sí mismas podrían tener un valor considerable al revelar las opiniones de la gente y (una observación relacionada) proveer claves acerca de sus probables acciones futuras. El gobierno puede servir a las personas mejor si sabe lo que están pensando, y no habrá de saberlo si las intimida al punto de que su boca permanezca sellada. Además, la libertad pública para criticar a los funcionarios gubernamentales disminuye los costos de agencia dentro del gobierno al sacar a la luz los atracos que cometen. De ahí que aplicar un castigo a las filtraciones de material clasificado como confidencial sea problemático, a menos que la clasificación esté basada en una necesidad de seguridad nacional de carácter urgente. Con frecuencia no lo está; con frecuencia el material se clasifica como confidencial a fin de cubrir las fechorías oficiales o exagerar la importancia de la información reunida por la agencia.

Una razón que tienen los conservadores para apoyar un derecho constitucional a quemar la bandera estadunidense es que probablemente reduce en vez de aumentar la incidencia de actos de incineración de esa especie. La quema de la bandera transmite un indicio más débil de la intensidad que caracteriza a las concepciones políticas del incinerador si éste no puede ser castigado por su acción. Entre más costoso sea un indicio, más enérgica será probablemente la creencia que lo indujo; tal es la razón de que los martirios cristianos fuesen un método efectivo para generar proselitismo.

XXIX.3. Restricciones previas, restricciones al punto de vista y el derecho de los foros

Históricamente, la *bête noire* de los defensores de la libertad de expresión ha sido la "restricción previa", es decir, el empleo de la licencia u otros métodos de censura para impedir que una obra sea publicada en primer lugar, por oposición al castigo del autor o el editor en un proceso penal (o quizás un juicio civil por difamación) después del hecho. Hay tres diferencias importantes

entre la regulación *ex ante* y *ex post* de la expresión. La primera es que las personas que hacen la regulación pueden ser diferentes. En el caso de la regulación *ex ante* habrá probablemente funcionarios administrativos (miembros de una junta de censores), quienes pueden haber sido atraídos a su trabajo por una creencia en la censura y en todo caso posiblemente conciben su misión como la de resolver dudas para al fin desautorizar una publicación y tal vez operen sin salvaguardas procedimentales. En el caso de la regulación *ex post* serán jueces y jurados, representativos de una extracción más o menos aleatoria de la población por lo que toca a las actitudes hacia la censura. No obstante, hoy en día la forma más común de regulación *ex ante* de la expresión es un mandato preliminar emitido por un juez, quien neutraliza las principales objeciones a las restricciones previas.

Segundo, los procedimientos pueden ser diferentes, aun cuando el regulador *ex ante* sea un tribunal en vez de una junta de censura. En particular, el castigo penal requiere que el fiscal soporte una carga de prueba pesada, y esto le resultará particularmente difícil de hacer en áreas tales como la obscenidad donde los "hechos" son escurridizos. Empero, esta objeción a las restricciones previas es superficial. Un castigo penal puede con facilidad aplicarse de manera lo bastante severa para tener el mismo efecto, a través de la disuasión, que una prohibición. Y para un editor adverso al riesgo sería preferible que le dijeran de antemano que no puede publicar algo en lugar de aventurarse a llevarlo a la imprenta y ser arrojado a prisión por el problema que provocó.

Tercero, el carácter de bien público de la información crea situaciones donde la regulación *ex ante* podría ser necesaria para evitar un daño irreparable. Supongamos que una revista está a punto de imprimir un artículo que explicará al terrorista más lerdo un medio viable para ensamblar una bomba atómica a bajo costo. Aquí la justificación de una prohibición preliminar u otro remedio *ex ante* es una aplicación directa de la fórmula presentada antes, en XXII.3; habiendo fallado la disuasión, el auxilio de la prohibición ordenada es el único remedio y sus beneficios están llamados a exceder sus costos.

Otra cuestión de técnica de cumplimiento de la ley es la elección entre centrarse en la expresión o discurso particularmente maligno e insistir en que la regulación del discurso sea neutral con respecto al punto de vista. La insistencia de los tribunales en esta última se relaciona con la política que se encuentra detrás de la cláusula de la protección igual de la Decimocuarta Enmienda. Entre más general sea una ley, más inadecuada será como una herramienta de abuso. Es más difícil de prevenir que los nazis marchen por las vecindades judías mediante una ley que prohíba todas las marchas que por medio de una ley que prohíba simplemente que los nazis marchen, porque otros grupos desean marchar y ejercerán su influencia política en contra de la ley más amplia. Pero esta técnica de protección de la libertad de expresión peca a menudo de resultar generalista y ser aplicada con poco esfuerzo. Los costos de diferentes medios de expresión varían, a menudo sistemáticamente, entre los gru-

pos, de modo que la prohibición de un medio particular podría alterar el equilibrio en el mercado de las ideas. Una ley contra los camiones de sonido o la venta de puerta en puerta afectará a los testigos de Jehová mucho más que a los unitarios. En cambio, el daño puede ser inseparable del punto de vista. Supongamos que el único tipo de pornografía que incita a los hombres a la violación es el de la pornografía que presenta mujeres bajo una luz degradante; de nada serviría prohibir la pornografía que careciera de este "punto de vista" particular.

En lo referente a la regulación del mercado de ideas y opiniones que es específica del sitio, los tribunales distinguen entre lo que llaman foros públicos tradicionales, foros públicos designados, foros públicos limitados y foros no públicos. La primera clase consiste principalmente en calles públicas, aceras y parques, que han estado tradicionalmente disponibles para la reunión y las manifestaciones públicas. La segunda consiste en sitios de propiedad pública que, aunque no se dedican tradicionalmente a propósitos de expresión, el gobierno ha decidido abrir para tales propósitos. La tercera (a menudo tratada en los juicios como parte de la segunda) consiste en sitios públicos configurados para un tipo de expresión específica, por ejemplo, un teatro de propiedad pública. La cuarta incluye todos los demás inmuebles públicos, algunos de los cuales podrían ser adecuados para la actividad expresiva (por ejemplo, las calles y las aceras de una base militar o la explanada de un aeropuerto), pero ninguno de los cuales está destinado a tales actividades. En las primera y segunda categorías, el gobierno puede regular sólo el tiempo, lugar y modo del discurso; en la tercera categoría puede confinar el tipo de discurso al tipo para el cual está diseñada la instalación; en la cuarta puede limitar al discurso como lo desee, siempre que mantenga la neutralidad entre puntos de vista rivales (esta reserva se aplica también a las otras tres categorías, por supuesto).

Estas distinciones tienen sentido económico en términos generales. Los foros públicos tradicionales y los designados pueden emplearse para actividades expresivas sin gran costo, siempre que se impongan restricciones para impedir el hacinamiento, y de este modo están permitidos, con lo que se alcanza un equilibrio entre costos y beneficios, tal como ocurre en otros ámbitos del derecho. El foro de propósito limitado, en contraste, no sería viable si no pudiera ser limitado a la función que tiene asignada. Imaginemos lo que le ocurriría a un teatro o una sala de conciertos (el Centro John F. Kennedy para las Artes Escénicas) si la administración tuviera que permitir que se usara para manifestaciones, reuniones políticas, piquetes y actividades similares. Éste es un caso en el que la restricción del discurso, o de un tipo de discurso, puede promover efectivamente la libertad de expresión, o una forma de expresión distinta y a la vez más valiosa. Hay muchos ejemplos al respecto.[10] Uno

[10] Véase Albert Breton y Ronald Wintrobe, "Freedom of Speech *vs.* Efficient Regulation in Markets for Ideas", 17 *J. Econ. Behav. & Organization,* 217 (1992).

es el concepto bifronte de "libertad académica". Remite simultáneamente a la libertad de expresión de un profesor y a "la autoridad de una universidad para administrar una comunidad académica y evaluar la enseñanza y el atesoramiento del saber encontrándose libre de interferencias por parte de otras unidades de gobierno, incluidos los tribunales".[11] El discurso, al igual que otros "productos", requiere una estructura procedimental, la más obvia de las cuales consiste en poner límites a la costumbre de interrumpir a otro hablante.

Otro ejemplo, si bien más debatible, de restricción del discurso impuesta para promoverlo es la ley del *copyright*, que al limitar el copiado incrementa los incentivos financieros para crear propiedad intelectual. Pero, al mismo tiempo, la protección del *copyright* reduce los incentivos financieros para mejorar las obras de carácter literario o de otra índole expresiva. Las extraordinarias mejoras que escritores como Shakespeare y Milton añadieron a las obras de las que tomaron en préstamo ciertos elementos, si se realizaran hoy en día, serían infracciones al *copyright*.[12]

XXIX.4. Análisis costo-beneficio de la libertad de expresión: el enfoque por categorías

Las grandes dificultades implicadas en cuantificar los costos y los beneficios de la libertad de expresión en un método de caso por caso han dado lugar a sugerencias para reconocer una jerarquía de categorías del discurso donde, por ejemplo, el discurso político y el científico reciban la máxima protección de acuerdo con el supuesto de que son los más valiosos desde un punto de vista social o de que (en el caso del discurso político) son más fundamentales para los intereses que animaron a los artífices y ratificadores de la Primera Enmienda. Con este enfoque, la publicidad comercial recibiría menos protección, junto con el arte y el entretenimiento, incluida la pornografía, mientras que las amenazas de cometer un crimen y la pornografía no recibirían ninguna protección.

El enfoque por categorías opera muy bien con la última categoría. Si el *único* propósito y el efecto probable de un discurso es crear una actividad indudablemente delictiva, puede demostrarse que tal discurso no tiene valor social y sólo se discutirá si los costos del discurso superan los costos de su prohibición. Lo que dificultó los casos comunistas, tales como Dennis, era que el Partido Comunista de los Estados Unidos era el agente conspirador de un gobierno extranjero hostil a la vez que una fuente de ideas interesantes

[11] Hosty *vs.* Carter, 412 F.3d, 731, 736 (7th Circ., 2005) (*en banc*); véase también Grutter *vs.* Bollinger, 539 U.S., 306, 328-330 (2003).

[12] Véase Richard A. Posner, *Law and Literature,* 525-529 (2003). Por supuesto, esta aserción es válida a condición de que las obras de las que tomaron algo en préstamo aún tengan vigentes sus derechos de autor.

acerca de la economía, las clases sociales, el racismo, la política exterior y otros fenómenos sociales importantes.

Sin embargo, dentro del conjunto general de los tipos de discurso que confieren algunos beneficios legítimos, el enfoque de la jerarquización (o léxico) falla porque confunde los beneficios totales con los marginales (véase antes XXV.4). Desde un punto de vista social general, podría ser peor prohibir todos los discursos políticos que toda manifestación artística; pero ésa no es la elección afrontada por una junta de censores u otro tribunal a quien se pida que prohíba un discurso particular, ya se trate de un anuncio comercial, un programa de televisión violento, una novela lesbiana gráfica o un espectáculo de club nocturno con mujeres desnudas. Aunque el discurso político en conjunto sea socialmente más valioso que la novela, un panfleto en defensa del genocidio podría tener menor valor social que una novela de Henry Miller que sea claramente no política; esto es posible aunque algo del placer obtenido por los lectores de la lectura de una novela de Miller derive de sus elementos pornográficos. No es una respuesta decir que no puede confiarse en que el gobierno permita la crítica de sí mismo; es probable que reprima cualquier discurso que sea radicalmente impopular, ya sea político, religioso, comercial o estético, con consecuencias no relacionadas con ninguna jerarquía de valores de la expresión.

Si no es viable estimar los beneficios del discurso por los métodos del derecho, un enfoque de costo-beneficio deberá concentrarse en los costos del discurso. Tales costos son a veces prohibitivos; un asesinato político puede ser en ocasiones un método efectivo para influir en la opinión política —más todavía que la quema de la bandera—, pero se le juzga a tal punto costoso que, en vez de ser privilegiado, es penalizado. En el otro extremo, los costos sociales del discurso son claramente nulos o aun negativos, como en Posadas de Puerto Rico Associates *vs.* Tourism Company of Puerto Rico,[13] donde la Corte Suprema sostuvo, con todo, que Puerto Rico, dado que podría haber prohibido constitucionalmente toda clase de juegos de azar, no había violado la Primera Enmienda al prohibir en su territorio (pero no fuera de allí) la publicidad de los juegos de casino legalizados por el gobierno de Puerto Rico. La decisión de la Corte no tiene ningún sentido económico.[14] El propósito ostensible de prohibir la publicidad era que se redujera la atracción de los juegos de azar para los residentes de Puerto Rico[15] (un propósito que no era incongruente con la legalización de los juegos de casino, porque eso podría haberse hecho simplemente en reconocimiento de la inutilidad de tratar de suprimirlos). La prohibición podría haber tenido algún efecto en ese sentido. Pero al mismo tiempo disminuía los costos de publicidad de los casinos. Esta disminución

[13] 478 U.S., 328 (1985).

[14] Véase Fred S. McChesney, "De-*Bates* and Re-*Bates:* The Supreme Court's Latest Commercial Speech Cases", 5 *Supreme Ct. Econ. Rev.,* 81, 102-105 (1997).

[15] Al permitir que los casinos hicieran publicidad fuera del estado, ¡Puerto Rico demostró que no le importaba nada la corrupción de los no residentes!

debiera conducirlos a disminuir sus precios, lo que a su vez volvería los juegos de azar más atractivos que antes de la prohibición. Por lo tanto, era probable que la prohibición privara a los consumidores de una información valiosa sin disminuir ninguno de los efectos colaterales indeseables de la adicción al juego (como la quiebra, la miseria, la malversación o el suicidio); es posible que incluso los hubiese incrementado.

XXIX.5. La regulación de la radiodifusión

Una de las excepciones a la libertad de expresión merece atención especial porque es una aplicación explícita de la teoría económica. Tal es el principio, ilustrado por la decisión de la Corte Suprema en Red Lion,[16] de que el gobierno puede regular el contenido de las transmisiones, por ejemplo, para asegurar la equidad, el equilibrio y la diversidad. La regulación no es una interferencia tan agresiva con el mercado de las ideas como el castigo por la expresión de ideas particulares, pero sí disminuye la libertad del transmisor para decidir cuáles ideas transmitirá, así como su incentivo para ventilar cuestiones controvertibles. Solamente si estando en el aire ventila asuntos controvertidos queda sujetado a la doctrina de la equidad. Esta sujeción debe imponerle un costo. Siempre es libre de poner a discusión ambas caras de un asunto, de manera que dicha doctrina sólo impone su rigor cuando él prefiere ver un solo lado de la moneda.

Se ha sostenido que una regulación similar, que aplica la "doctrina de la equidad" a los periódicos, viola la Primera Enmienda.[17] Según la Corte, la diferencia reside en que los radiodifusores, al revés de lo que ocurre con los propietarios de periódicos, tienen un poder monopólico. Dado que dos radiodifusores no podrían transmitir en la misma frecuencia en la misma área sin crear una interferencia intolerable, una licencia de la Comisión Federal de Comunicaciones (FCC) para el uso de una frecuencia particular en un área particular confiere un monopolio de la frecuencia que no tiene contrapartida en otros medios de expresión.

Éste es un absurdo desde el ángulo económico. Es cierto que sólo una frecuencia puede usarse en el mismo lugar al mismo tiempo, pero el resultado no es necesariamente un monopolio, ya que diferentes frecuencias son, dentro de cierto intervalo muy amplio, sustitutos perfectos entre sí. La FCC licencia generalmente a más de una estación de televisión en cada mercado. La mayoría de los mercados tienen por lo menos 3 o 4 estaciones, y algunos tienen de 9 a 11. Esto es invariablemente más que el número de periódicos en el mismo mercado. Y éstas son sólo estaciones que están al aire; hay ahora gran número

[16] Red Lion Broadcasting Co. *vs.* FCC, 395 U.S., 367 (1969). Véase también National Broadcasting Co. *vs.* United States, 319 U.S., 190, 226 (1943).

[17] Miami Herald Publishing Co. *vs.* Tornillo, 418 U.S., 241 (1974).

de canales de televisión por cable independientes. Los suscriptores a la televisión por cable o satélite a menudo tienen acceso a cientos de canales.

El hecho de que el espectro electromagnético esté limitado no lo distingue de otros recursos. Los insumos empleados en los métodos alternativos para la comunicación de ideas son también limitados. De hecho, los periódicos padecen escaseces más agudas que, aunadas a la presión de la competencia de sustitutos, han generado un grado de monopolio local mucho mayor que en la radiodifusión. La reducción de la demanda de periódicos, debida al ascenso de la televisión pero mucho más a la Internet, ha acentuado las condiciones de monopolio natural de la producción de periódicos; como sabemos, estas condiciones son más importantes entre menor sea la demanda del producto del mercado (véase antes XIV.1). Muchos de los costos de un periódico son fijos: en particular el costo de la investigación y la redacción de los artículos y reportajes, y el costo de la composición y la tipografía para ellos y para los anuncios. El costo marginal —el costo de la impresión de otro ejemplar del periódico— es pequeño. Por lo tanto, a menos que haya una gran demanda, es improbable que exista espacio para más de un periódico por comunidad.

Supongamos que un radiodifusor tiene un monopolio efectivo del mercado en que opera. ¿Cuál impacto sería de esperarse sobre la difusión de las noticias y las opiniones? El radiodifusor podría limitar la cantidad de tiempo que trabaja en el aire, y esto disminuiría la distribución de ideas entre quienes están en el mercado. Pero la doctrina de la equidad no está calculada para aumentar la producción de ideas por el radiodifusor. Por el contrario, lo castiga por presentar ideas controvertidas al exigirle que presente todos los lados de una controversia, de modo que induce la presentación de ideas no controvertidas.

¿Podría el radiodifusor monopólico distorsionar las noticias y suprimir las opiniones con las que no esté de acuerdo, tratando de convertir a los televidentes a sus propias opiniones? Todo radiodifusor podría hacer esto; la importancia del monopolio reside en que el costo para el radiotransmisor es menor que en un mercado competitivo y el daño para los televidentes es mayor. Es menor porque sus televidentes no tienen buenos sustitutos y porque el control de su fuente de información por parte del radiodifusor hará que sea más difícil para ellos descubrir que están siendo engañados. El costo para ellos es mayor que en el caso competitivo donde la parcialidad de un radiodifusor no les negaría el acceso a otras opiniones. Pero las ganancias del radiodifusor monopólico serán menores si introduce sus preferencias personales en la decisión de lo que habrá de transmitirse. Y los administradores de una gran corporación pública tendrán dificultad para encontrar cuestiones donde una postura parcial no aliene a gran número de los accionistas.

Otra posibilidad es que la estación monopólica altere la combinación de las ideas transmitidas sólo como un subproducto del esfuerzo para maximizar su ingreso pecuniario; vimos en el capítulo IX que un monopolista podría pro-

ducir un bien de calidad diferente de la del bien de una empresa competitiva (véase antes IX.3). No obstante, la estación monopólica podría producir fácilmente una emisión de calidad mayor, antes que menor. Supongamos que hay dos estaciones de televisión en un mercado dado, cada una de ellas propiedad de una empresa diferente, y que cada una compite vigorosamente con la otra. Es probable que cada una trate de atraer la audiencia más grande posible a fin de maximizar su ingreso de publicidad. Entonces será de esperarse que las estaciones transmitan materiales muy similares, como ocurre con los dos partidos principales de un sistema político bipartidista que son de ordinario centristas. Sin embargo, un solo propietario de ambas estaciones podría presentar materiales muy diferentes en las dos estaciones a fin de maximizar la audiencia total de ambas.

Al momento de redactar estas líneas, los dos grandes partidos políticos de los Estados Unidos son menos centristas que antes, y esto ha dado pábulo a una cobertura de los medios políticamente polarizada. La mayoría de las personas no desean quedar expuestas a la presentación de ideas y opiniones que cuestionan sus juicios políticos o religiosos o morales previos. No tienen por qué encontrarse una vez más en una situación así. Esto produce la paradoja de que, entre más competitivo resulta el mercado de las ideas y opiniones, menor es el valor que produce; se convierte en una cámara de ecos en lugar de ser un estímulo para el pensamiento. La competencia no siempre es buena, y el monopolio, aun cuando no esté regulado, no siempre es malo.

XXIX.6. La publicidad falsa y la relación entre los derechos políticos y los económicos

Si el vendedor de un fármaco anuncia que puede curar la artritis, la Comisión Federal de Comercio (FTC) puede prohibir la publicidad si se demuestra que es probablemente falsa. Si la misma aseveración se hiciera en un libro, estaría claro que la Primera Enmienda prohibiría a la FTC tal precepto negativo, por lo menos si el autor no fuese el vendedor del fármaco.[18] Es curioso el tratamiento diferente de los dos casos. Dado que las afirmaciones son idénticas, la diferencia no puede ser que la publicidad sea obviamente cierta u obviamente falsa; está claro que a menudo no es una cosa ni la otra. Tampoco puede ser el factor distintivo en la presencia de un motivo comercial en un caso pero no en el otro. El autor de un libro sobre la salud tendrá un incentivo para hacer afirmaciones falsas si cree que aumentará su ingreso por la venta del libro, a menos que piense que el exponerse lo dañaría más; pero el riesgo de la exposición es el mismo para el vendedor del fármaco. Los candidatos políticos tienen un incentivo para hacer afirmaciones falsas si creen que tales

[18] *Cf.* Rodale Press, Inc. *vs.* Federal Trade Commission, 407 F.2d, 1252 (D.C. Cir., 1968).

afirmaciones los acercarán al poder, y lo mismo ocurre con los profesores y la reputación académica.

Así, ¿cómo explicar el tratamiento diferente del discurso que no es comercial ni político (principalmente la expresión artística), al que se otorga una protección casi completa de acuerdo con la Primera Enmienda,[19] del tratamiento otorgado al discurso comercial, al que se le da una protección menor (aunque no nula)? La Corte Suprema ha empezado ahora a cerrar la brecha entre el tratamiento del discurso comercial y el no comercial de acuerdo con la Primera Enmienda, como en su decisión de invalidar una ley que prohibía a las farmacéuticas anunciar los precios que cobran por los medicamentos con receta.[20] La Corte pensaba que la restricción era poco razonable y también especialmente onerosa para los pobres. Por ello, ahora todo caso de publicidad falsa que se plantee ante la FTC, como todo caso de difamación, plantea una cuestión potencial de la Primera Enmienda,[21] aunque tienda a resolverse a favor de la Comisión Federal de Comercio.

Pero la diferencia en el trato concedido a la expresión comercial y la no comercial persiste. Se relaciona con una dicotomía más amplia: aquella entre la protección de la competencia en materia de bienes y la competencia en materia de ideas, que es característica del pensamiento moderno.[22] Los clásicos liberales, como John Stuart Mill, creían tanto en la libertad económica como en la intelectual. En la era moderna, la preferencia de la Corte Suprema por esta última libertad puede reflejar la especial importancia que tienen los derechos políticos a ojos de los abogados, jueces y eruditos en materia constitucional: se trata de personas con un fuerte interés (en ambos sentidos de la palabra) en los escenarios públicos y políticos de la acción. Aun así, los derechos económicos poseen la misma importancia para una porción más extensa, si bien menos capaz de expresarse con elocuencia, de la población. Y el gobierno con frecuencia infringe esos derechos. La legislación puede limitar la elección de la ocupación, transferir riqueza de los consumidores a los accionistas e impedir que las personas consigan los servicios que desean y por los que están dispuestas a pagar.

[19] Con la excepción de la obscenidad.

[20] Virginia State Board of Pharmacy *vs.* Virginia Citizens Consumer Council, 425 U.S., 748 (1976). Véanse algunas pruebas de que tales leyes elevan los precios en Lee Benham, "The Effect of Advertising on the Price of Eyeglasses", 15 *J. Law & Econ.,* 337 (1972). Algunos casos subsecuentes de la libertad del discurso comercial son los de Sorrell *vs.* IMS Health, 131 S. Ct., 2653 (2011); Thompson *vs.* Western States Medical Center, 535 U.S., 357 (2002); Bolger *vs.* Youngs Drug Products Corp., 463 U.S., 60 (1983) y, por supuesto, Posadas de Puerto Rico Assocs. *vs.* Tourism Co. of Puerto Rico, discutido antes en XXIX.4.

[21] Como en Encyclopaedia Britannica, Inc. *vs.* FTC, 605 F.2d, 964, 972-973 (7th Cir., 1979).

[22] Véase Aaron Director, "The Parity of the Economic Market Place", 7 *J. Law & Econ.,* 1 (1964); Ronald H. Coase, "The Market for Goods and the Market for Ideas", 64 *Am. Econ. Rev. Papers & Proceedings*, 384 (mayo de 1974); Coase, "Advertising and Free Speech", 6 *J. Legal Stud.,* 1 (1977).

Ahora bien, ¿hemos pasado por alto una diferencia fundamental entre la libertad de expresión y la libertad de comercio: que la primera es necesaria para impedir lo que señalamos antes que era la forma más peligrosa del monopolio, mientras que la segunda no lo es? Sin embargo, es posible que la libertad económica y la política no puedan separarse tan claramente. El disentimiento político requiere recursos financieros. En una sociedad donde el gobierno controlara toda la actividad económica —donde el papel estuviera racionado, la impresión estuviera sujeta a licencias y el Estado fuera, directa o indirectamente, el empleador principal—, sería muy difícil organizar y financiar la actividad política en oposición al gobierno. En los días de gloria del senador Joseph McCarthy, las personas de quienes se creía que simpatizaban con el comunismo eran excluidas del empleo gubernamental, incluso de empleos no confidenciales. Estas personas no se morían de hambre. Encontraban empleos en el sector privado y algunas volvían a actuar en la política finalmente. Los costos del disentimiento habrían sido mayores si el gobierno hubiese sido el único empleador, de modo que las consecuencias de tener opiniones impopulares podrían ser la negación de toda oportunidad de obtener un medio de vida.

Los derechos políticos y los económicos convergen en otro sentido. La tendencia de la adjudicación constitucional es hacia reconocer los derechos especiales de la protección constitucional de grupos diferentes de las minorías raciales, religiosas y políticas tradicionales, como el de las mujeres. Sin embargo, sus intereses son con frecuencia idénticos al interés público más amplio en la libertad económica. La invalidación de las restricciones impuestas a las elecciones ocupacionales de las mujeres promueve la eficiencia.

Empero, hay dos argumentos económicos para otorgar al discurso comercial una protección de la Primera Enmienda menor que la del discurso político, artístico y científico… y quizá para no otorgar ninguna protección. Primero, la veracidad de las aseveraciones típicas que se hacen en la publicidad puede estimarse por los tribunales con mayor certeza y a un costo menor que el valor de las ideas políticas, artísticas o científicas. Segundo, el orador comercial espera con frecuencia recuperar el valor económico pleno de su discurso con la forma de precios mayores de los productos o de una producción mayor. Es decir, pocos de los beneficios generados por la publicidad y otros discursos comerciales son externos, de modo que resulta dudoso que el discurso comercial debiera obtener una protección constitucional mayor que la actividad comercial en general. En cambio, a menudo una parte grande de los beneficios de las ideas políticas, artísticas y científicas son externos al creador, mientras que los costos se concentran en él. En efecto, la Primera Enmienda externaliza algunos de esos costos (el daño causado a otros por el discurso), tratando de revertir el equilibrio. Ya hemos observado que ésta es una técnica empleada por la ley (véase antes VI.4).

XXIX.7. El subsidio del discurso

Algunas intervenciones públicas en el mercado del discurso no asumen la forma de la supresión del discurso reprobado sino la del subsidio del discurso contrario, ya se trate del consejo de la conducción segura, el sexo seguro o los valores patrióticos, o de proveer apoyo financiero para las campañas políticas. Ésta es una interferencia gubernamental en el mercado de las ideas y las opiniones, e implica la coerción de los contribuyentes que no están de acuerdo con la posición que el subsidio apoya y no desean que su dinero se emplee para propagarla. Pero el subsidio y la prohibición no son simétricos en sus efectos. Un gasto dado en dólares por el gobierno tendrá un efecto mayor si se destina a cumplir una prohibición del discurso que si se destina al financiamiento del discurso contrario. La publicidad es muy cara y probablemente ineficaz cuando aconseja el comportamiento contrario al del interés propio de la audiencia. La segunda observación muestra la inutilidad del discurso contrario dirigido, por ejemplo, contra la pornografía o contra la violencia en el cine; la gente que disfruta estas cosas no se sentirá conmovida por los anuncios que le dicen que no las disfrute. La publicidad contra el tabaco que destaca los efectos de fumar sobre la salud será más probablemente eficaz, porque apela al interés propio de la audiencia, que la publicidad destinada sólo a hacer que los consumidores de pornografía se sientan culpables o "sucios".

La publicidad contra el aborto es un caso intermedio. La audiencia está integrada por personas que tienen oportunidades de elegir: ya sean mujeres embarazadas, los padres de jóvenes adolescentes o las mujeres jóvenes o adultas que deciden tener sexo o no, y cuál tipo de anticonceptivos emplearán. Estas elecciones implican un entrelazamiento de consideraciones de interés propio con preocupaciones altruistas. Si las elecciones están finamente balanceadas, la propaganda pública podría inclinar la balanza en algunos casos. Sin embargo, comparemos el gasto gubernamental de 100 millones de dólares en una campaña de publicidad contra el aborto y la misma cantidad de dinero gastada en la persecución de quienes defienden el aborto. En el último caso, la defensa del aborto será empujada hasta un nivel muy bajo y, en consecuencia, bajará el número de abortos, aunque un efecto contrario parcial sería que resultarían más baratos porque las clínicas para abortar estarían economizando en la publicidad, como en el caso de la regulación de la publicidad de los casinos en Puerto Rico (véase antes XXIX.4). El efecto del gasto de 100 millones de dólares en publicidad para denunciar el aborto sería menor; porque la publicidad gubernamental contra el aborto sustituiría a los anuncios financiados por el sector privado, estos últimos disminuirían y el aumento neto de la publicidad contra el aborto podría ser escaso. Y los anuncios gubernamentales darían a su audiencia una información que ya tenía, con publicidad o sin ella: una información balanceada, además, y en esa me-

dida contrarrestada, por la publicidad en favor de la elección. En suma, el enfoque del subsidio tendría, dólar por dólar, un impacto menor sobre el mercado de las ideas y opiniones que el enfoque de la regulación.

XXIX.8. La economía de la libertad de religión

La Primera Enmienda no protege sólo las libertades estrechamente relacionadas de expresión, petición, reunión y prensa, sino también la libertad de religión. La enmienda prohíbe al gobierno (inicialmente sólo al gobierno federal, pero, al igual que en el caso de la mayoría de las demás disposiciones de la Carta de Derechos de los Estados Unidos, ahora también a los gobiernos estatales y locales mediante la interpretación de la cláusula del debido proceso de la Decimocuarta Enmienda) "establecer" una religión o interferir con su libre ejercicio. De este modo, el gobierno debe mostrarse neutral ante las religiones pero, asimismo, de acuerdo con la interpretación actual de la Primera Enmienda, neutral ante la alternativa entre religión e irreligión. Consideremos ahora la economía de establecer ambas en su sentido convencional (una "iglesia establecida") y con una interpretación más amplia de la Corte Suprema.

Toda nación de Europa Occidental cuenta con una y, a veces, con más de una iglesia establecida (esto es, apoyada o subvencionada por el Estado). Uno podría esperar, por consiguiente, que los Estados Unidos se encontraran entre las naciones menos religiosas. En realidad es una nación más religiosa, según se mide por el porcentaje de estadunidenses que se identifican a sí mismos como religiosos, que cualquier nación europea aparte de Polonia e Irlanda.[23] El gran filósofo británico del siglo XVIII David Hume argumentaba con fundamentos esencialmente económicos que una iglesia establecida debilita en lugar de fortalecer la creencia religiosa, así como (un punto relacionado) que, en lugar de fomentar los conflictos religiosos (una inquietud que afectaba a los artífices de la Constitución), propicia la apatía religiosa. Hume pensaba que los representantes religiosos pagados por el gobierno actuarían a semejanza de otros servidores civiles —un grupo no conocido por su fanatismo— porque no tendrían ningún incentivo pecuniario para convertir a la gente o maximizar la asistencia al templo.

Un punto menos obvio es que el subsidio público de una iglesia en particular les creará a las demás iglesias más dificultades para competir. Como resultado, habrá menos variedad religiosa que en el caso en que la competencia se desarrollara sobre un campo de juego con reglas iguales para todos. Una reducción en la variedad de los productos (sin ninguna reducción en el costo) habrá de reducir la demanda del producto. Pero este punto es menos persuasivo que

[23] WIN-Gallup International, Global Index of Religiosity and Atheism, 6 de agosto, 2013, cuadro 1: Global Religiosity Index for 2012, www.wingia.com/web/files/news/14/file/14.pdf.

el elaborado por Hume. El subsidio puede estimular la demanda por la iglesia establecida al reducir el costo ajustado a la calidad de asistir a ella; supongamos que el subsidio se emplea para construir magníficas catedrales o contratar a destacados organistas y coros bien afinados. Además, si el subsidio provoca que los funcionarios de la iglesia establecida se vuelvan indolentes, esto puede contrarrestar su ventaja en costos y facilitar la competencia de otras sectas.

El argumento de Hume es quizá inaplicable a las formas atenuadas del establecimiento de que todas son posibles en una sociedad pluralmente religiosa como los Estados Unidos (por supuesto, este país puede ser plural en parte por la razón de Hume). Tal vez en ningún estado salvo Utah podría ser aprobada una ley que estableciera una secta religiosa particular aun cuando la Cláusula de Establecimiento nunca se hubiese juzgado aplicable a los estados. Una exhibición de los Diez Mandamientos en propiedad pública se halla muy alejada de un ministro pagado por el Estado en lo tocante al impacto que tiene el apoyo público a la religión sobre el proselitismo.

Sin embargo, ¿es preferible que la sociedad tenga más denominaciones religiosas o, como creía Hume, menos? Adam Smith creía que, entre más denominaciones hubiera y, por ende, menor en promedio fuese cada una de ellas, más eficaz sería la religión en la regulación de la moral, porque disminuiría el problema del "gorroneo". Probablemente estaba equivocado; es probable que, entre más denominaciones religiosas deban competir por adeptos, más fácil será que una persona de moral laxa encuentre una denominación complaciente para unirse a ella. Una religión monopólica puede ser severa porque tiene un mercado cautivo; las religiones competitivas no pueden darse el lujo de ser severas, así como tampoco los vendedores competitivos de bienes y servicios ordinarios pueden darse el lujo de tratar a sus clientes de un modo censor.

Casi todos los casos vinculados a la Cláusula de Establecimiento han implicado esfuerzos para "establecer" la religión en sectas generales (en contraposición a la incredulidad) en vez de hacerlo en sectas particulares. Esos esfuerzos han adoptado formas tales como la de reservar tiempo para la oración voluntaria en escuelas públicas, fomentando en éstas la instrucción escolar respecto al "diseño inteligente", suministrando fondos públicos para la educación secular en escuelas religiosas (principalmente católicas) o para la representación del Nacimiento, o, como sucedió en dos casos decididos en fecha reciente por la Corte Suprema, exhibiendo materiales religiosos en propiedad pública, por lo general sin costo para el público; resulta muy fácil obtener donaciones de tales materiales, como en el caso del monumento de los Diez Mandamientos otorgado al estado de Texas por la Orden Fraternal de las Águilas (FOE) por sugerencia de Cecil B. DeMille, el productor de la película *Los diez mandamientos*.[24]

[24] Van Orden *vs.* Perry, 545 U.S., 677 (2005). El otro caso es el de McCreary County *vs.* American Civil Liberties Union of Kentucky, 545 U.S., 844 (2005).

Algunos de estos esfuerzos se realizan a fin de violar la Cláusula de Establecimiento, otros no; no existe un patrón discernible o un estándar legal compacto. Desde una perspectiva económica, dada la premisa de la neutralidad gubernamental, la legalidad de tales "establecimientos" debería inducirnos a preguntar si el probable efecto buscado no es meramente el de equipararse a algún subsidio para actividades seculares de competencia. De acuerdo con este criterio, las numerosas decisiones de la Corte Suprema en torno a la Cláusula de Establecimiento son como un edredón hecho de puros retazos. Se permite que las instituciones religiosas queden exceptuadas del impuesto predial y, aun así, a semejanza de la excepción similar concedida a las instituciones educativas, la medida es justificable sólo si se puede pensar que esas instituciones confieren beneficios externos iguales o mayores al costo de los servicios públicos que reciben y que son pagados mediante los impuestos prediales que ellas no pagan. Nadie lo sabe. Y advirtamos que la excepción interfiere arbitrariamente con la competencia entre las sectas religiosas favoreciendo a las que se resuelven por las edificaciones más elaboradas.

Por otra parte, las decisiones que prohíben al gobierno pagar cualquier porción del costo de los colegios religiosos (un término utilizado en sentido amplio para referirse a cualquier escuela administrada por una organización religiosa, no solamente una escuela católica) discriminan en contra de la religión. Un colegio religioso produce dos resultados finales: educación secular y educación religiosa. Compite con las escuelas públicas y las privadas no adscritas a una secta, que producen sólo una educación secular. Si el gobierno sufragara por entero el costo de los colegios religiosos, estaría subsidiando a la religión, puesto que estaría desembolsando dinero para la educación religiosa. Pero si no sufraga un ápice del costo de las escuelas religiosas, está otorgando un subsidio a un competidor secular —las escuelas públicas— y por ese motivo está discriminando contra la religión. Así que debiera pagar por el componente secular que tiene el programa de las escuelas parroquiales. ¿Pero cómo se ha de cuantificar ese componente? La mayor parte de los costos que conlleva el dar acabado a los productos finales de naturaleza religiosa y secular en un colegio religioso son conjuntos —prácticamente todo el costo de las instalaciones escolares y del personal administrativo—, y recordamos por análisis previos (véase antes XIII.5) que la producción se maximiza cuando los costos conjuntos se asignan de manera inversa a la elasticidad de la demanda de bienes o servicios para cuya producción se incurre en ellos. Si la demanda de educación religiosa es más elástica que la demanda de educación secular por las escuelas parroquiales, la mayor parte de los costos conjuntos que representa la educación del colegio religioso deberían ser considerados como costos relativos al componente secular, que el gobierno estaría obligado a subsidiar con la finalidad de mantener la paridad con las escuelas públicas (a menos que se creyera que las escuelas públicas crean beneficios externos mucho mayores que los colegios religiosos).

Como consecuencia de la decisión adoptada por la Corte Suprema de defender los bonos escolares,[25] en la actualidad se le permite al gobierno cubrir la colegiatura o matrícula de la escuela parroquial, tal como se le ha permitido que lo haga con respecto a la colegiatura de las universidades católicas (y las de otras confesiones o denominaciones). El subsidio a la religión es de hecho más grande que en el caso en que el gobierno efectúa un pago por sólo los programas seculares de las escuelas religiosas, en virtud de que la matrícula cobrada por dichas escuelas cubre tanto los programas religiosos como los seculares que ofrece. ¿Pero se trata realmente de un subsidio a la religión? Las escuelas no religiosas son igualmente elegibles; la elección se deja en manos del consumidor.[26] Si acaso, el efecto general de los programas de bonos escolares puede consistir en debilitar a las escuelas religiosas al incrementar los recursos y la calidad de las escuelas seculares, tanto las públicas como las privadas. Las personas que se convierten al catolicismo a fin de que sus hijos acudan a una buena escuela pueden encontrarse con que los bonos en cuestión han estimulado la creación de más escuelas privadas, las cuales a su vez han ejercido más presión sobre las escuelas públicas a fin de que mejoren.[27]

Lecturas recomendadas

Bade, Karl, "Scylla, Charybdis and Adam Smith: An Economic Analysis of the Religion Clauses", 39 *DePaul L. Rev.,* 1235 (1990).

Breton, Albert y Ronald Wintrobe, "Freedom of Speech *vs.* Efficient Regulation in Markets for Ideas", 17 *J. Econ. Behav. & Organization,* 217 (1992).

Coase, Ronald H., "The Market for Goods and the Market for Ideas", 64 *Am. Econ. Rev. Papers & Proceedings,* 384 (mayo de 1974).

——— "Advertising and Free Speech", 6 *J. Legal Stud.,* 1 (1977).

Director, Aaron, "The Parity of the Economic Market Place", 7 *J. Law & Econ.,* 1 (1964).

Farber, Daniel A., "Free Speech without Romance: Public Choice and the First Amendment", 105 *Harv. L. Rev.,* 554 (1991).

Gans, Joshua S. y Scott Stern, "Is There a Market for Ideas?", 19 *Industrial & Corporate Change,* 805 (2010).

Hylton, Keith N., "Church and State: An Economic Analysis", 13 *Am. Law & Econ. Rev.,* 402 (2011).

McChesney, Fred S., "A Positive Regulatory Theory of the First Amendment", 20 *Conn. L. Rev.,* 355 (1988).

[25] Zelman *vs.* Simmons-Harris, 536 U.S., 639 (2002).

[26] Keith N. Hylton, "Church and State: An Economic Analysis", 13 *Am. Law & Econ. Rev.,* 402 (2011).

[27] Véase David Figlio y Cassandra M.D. Hart, "Does Competition Improve Public Schools? New Evidence from the Florida Tax-Credit", *Education Next,* invierno de 2011, 74, 76-78.

McConnell, Michael W., y Richard A. Posner, "An Economic Approach to Issues of Religious Freedom", 56 *U. Chi. L. Rev.,* 1 (1989).
Posner, Eric A., "The Legal Regulation of Religious Groups", 2 *Legal Theory,* 33 (1996).
Posner, Richard A., "Free Speech in an Economic Perspective", 20 *Suffolk U. L. Rev.,* 1 (1986).
West, E. G., "An Economic Analysis of the Law and Politics of Non-Public School 'Aid'", 19 *J. Law & Econ.,* 79 (1976).

Problemas

1. Blockbusting es la práctica por la que los agentes inmobiliarios tratan presuntamente de asustar a los propietarios blancos para que vendan sus casas a precios de ganga, diciéndoles que la vecindad se está volviendo negra. ¿Debiera estar esa práctica protegida por la Primera Enmienda? ¿Sería diferente si la información del vendedor al terrateniente fuese cierta?

2. Muchos tribunales han sostenido que una funeraria en un barrio residencial es una molestia. Véase "Comment", 20 *Syracuse L. Rev*., 45 (1968). ¿Es ésta una aplicación (o extensión) apropiada de la teoría económica del perjuicio (véase antes III.7)? Supongamos que pudiera demostrarse que la introducción de una librería pornográfica en una comunidad disminuye los valores de los inmuebles. ¿Debiera considerarse esa librería como un perjuicio?

3. En términos del análisis económico presentado en este capítulo, ¿debería sostenerse que una ley que prohibiera toda publicidad de cigarros en radio y televisión viola la Primera Enmienda? *Cf.* Capital Broadcasting Co. *vs.* Mitchell, 333 F. Supp., 582 (D.D.C., 1971), *aff'd*, 405 U.S., 1000 (1972).

4. ¿Hay algún argumento económico para regular la obscenidad más restrictivamente en la televisión o la radio que en los medios impresos?

5. Los defensores de la democracia deliberativa han propuesto a veces que el gobierno abra más sus propiedades de defensa. Véase, por ejemplo, Joshua Cohen, "Freedom of Expression", 22 *Philo. & Pub. Affairs,* 207, 247 (1993). Supongamos que los tribunales aceptaran la propuesta y decidieran que los aeropuertos y las escuelas debieran ser considerados foros públicos disponibles para todas las clases de defensas: marchas y otras manifestaciones, oratoria de banqueta, peticiones enérgicas. Elabore el argumento económico que demuestra que la propuesta reduciría más probablemente, en lugar de aumentar, la cantidad de la defensa.

6. ¿Prohibiría la Primera Enmienda un estatuto que declarara que constituye una nociva invasión de la privacidad revelar las relaciones sexuales extramaritales de un funcionario público consensuadas con otro adulto??

7. ¿Debieran tener los blogs los mismos derechos de la Primera Enmienda que los periódicos?

XXX. CATEOS, CONFISCACIONES E INTERROGATORIOS

XXX.1. Revisión del derecho a la privacidad

La Cuarta Enmienda garantiza el derecho de los ciudadanos a sentirse seguros en sus personas, domicilios, papeles y pertenencias contra los cateos y confiscaciones irrazonables, y también prohíbe las órdenes de detención generales (es decir, las órdenes de detención que no describan con particularidad la persona o cosa que será confiscada o cateada).[1] La cláusula de autoincriminación de la Quinta Enmienda prohíbe que el gobierno obligue a las personas a acusarse a sí mismas de un delito. Este capítulo discute los problemas económicos planteados por estas garantías constitucionales.

Como mínimo, podría pensarse que la Cuarta Enmienda protege a las personas contra los allanamientos a la persona o la propiedad por los funcionarios federales (y también a los estatales desde que se consideró que la enmienda es aplicable a los estados de acuerdo con la cláusula del debido proceso de la Decimocuarta Enmienda). Esto sería la protección de la privacidad en un sentido importante aunque limitado. Si la policía llega a la casa de usted e interrumpe lo que está haciendo para catearla, hay una invasión de la privacidad en un sentido muy semejante a lo que ocurre cuando una solicitud telefónica no deseada constituye una invasión de la privacidad, aunque la policía no obtenga ninguna información acerca de usted. Si la policía interviene su teléfono, hay una invasión de la privacidad en un sentido diferente. No hay ninguna interrupción —ninguna interferencia con su soledad o concentración—, pero hay una invasión de la privacidad considerada como ámbito de la confidencialidad. Aunque en el capítulo XI cuestionamos si vale la pena proteger la privacidad en este sentido cuando no están implicados los secretos comerciales, hay un argumento a favor de por lo menos una confidencialidad limitada de las comunicaciones telefónicas. Sin ella, las personas incurrirán en costos para cambiar a métodos de comunicación más seguros pero menos eficientes. Ésta no es una pérdida —es una ganancia— si las comunicaciones son ilegales; pero si son legales hay una pérdida, y esto provee un argumento económico a favor de la decisión de Katz,[2] donde se sostuvo que la intervención telefónica era una forma de confiscación dentro del significado de la Cuarta Enmienda y, por lo tanto, ilegal si es irrazonable (lo que significa que los costos exceden a los beneficios, como vamos a ver enseguida).

[1] Véase Wayne R. LaFave, *Search and Seizure: A Treatise on the Fourth Amendment,* vols. 1-6 (4ª ed., 2004).

[2] Katz *vs.* United States, 389 U.S., 347 (1967).

El hecho de que los cateos y las confiscaciones (incluidos los arrestos) impongan costos sociales no es, por supuesto, un argumento a favor de su prohibición. Pero es un argumento en favor de su regulación, de modo que la policía no realice cateos cuando los costos sociales excedan a los beneficios sociales. La palabra clave en la Cuarta Enmienda en lo tocante a la formulación de un estándar legal es la de "irrazonable", y esto debería invitarnos a emplear la Fórmula de Hand como guía para la indagación. Un cateo (o una confiscación) es razonable si el costo del cateo en términos de la afectación de la privacidad (*B*) es menor a la probabilidad (*P*) de que sin el cateo no pueda condenarse al destinatario del cateo o, de otro modo, volverlo inofensivo (el significado de "inofensivo" en esta frase de la cláusula se despejará en la última sección de este capítulo), multiplicada por la pérdida social (*L*) si elude el castigo. *P* tiene dos componentes: la probabilidad de que el cateo produzca algo de valor para la policía (la causa probable), y la probabilidad de que ese algo sea esencial para la condena. Por lo tanto, el valor del cateo es menor si la misma prueba pudiera obtenerse sin el cateo (así que presumiblemente con una *B* menor). Entre más invasivo sea el cateo y, por ende, mayor sea *B,* más esencial deberá ser la prueba buscada (*P*), o más grave deberá ser el delito investigado (*L*) para contrarrestar a *B*.

Los tribunales parecen en general conscientes de estos factores. Un cateo mínimamente invasivo (es decir, una *B* baja) —una parada y cacheo o una inspección mediante palmadas (una parada de Terry, como se la denomina, evocando el caso en el que la Corte Suprema determinó que semejante restricción era constitucional siempre y cuando hubiera sospechas razonables de actividad delictuosa)— es permisible sobre la base de una *P* menor a la de un cateo de la casa o un arresto. Entre más bajos sean los costos del cateo, menores serán los beneficios anticipados que deban ponerse a la vista para justificarlo. El caso limitante lo ilustra la decisión de la Corte Suprema en Illinois *vs*. Lidster,[3] que defendió, sobre el desafío impuesto por la Cuarta Enmienda, la razón de ser de un bloqueo al camino que la policía había instalado para detener a los automóviles para tratar de obtener información de los conductores acerca de un accidente reciente en el que el involucrado se dio a la fuga. No sólo no existían sospechas por parte de un individuo sobre las personas detenidas: ni siquiera había una sospecha colectiva. El propósito del bloqueo no era detener al automovilista que se dio a la fuga, sino recabar información que pudiera conducir a su aprehensión en otro lugar. Aun así, la Corte determinó que el bloqueo al camino no violaba la Cuarta Enmienda. Un factor en el que hizo hincapié era que obligar a los conductores a parar el auto significaba una invasión de menor escala que la de un arresto o un cateo convencional. Adviértase asimismo que el carácter *indiscriminado* de un bloqueo al camino o de alguna otra forma de lanzar una red para atraer para

[3] 540 U.S., 419 (2004).

sí a unos culpables es de hecho uno de sus rasgos más redimibles. A diferencia del perfil étnico, no discrimina entre los grupos, a menos que se utilice de manera desproporcionada en barrios claramente identificables desde el ángulo racial o étnico.

Lidster, uno de los conductores forzados a hacer parada ante el bloqueo, fue arrestado porque la policía se percató de que estaba borracho. Ésta fue una aplicación de la regla de "a simple vista"; si la policía emprende un cateo o una confiscación por una razón sólida y en el curso de sus operaciones descubre pruebas no anticipadas de un delito, puede utilizarlas sin necesidad de mostrar la probable causa o la sospecha razonable (o incluso la mínima). En términos económicos, la razón es que el costo incremental (en la invasión de la privacidad) para la persona objeto de cateo o confiscación es de cero.

Lidster constituye un importante precedente debido a que practica una bifurcación entre el cateo y la sospecha. Permite que se lleve a cabo la vigilancia que invade la libertad y la privacidad a causa de la importancia de la información buscada, aun cuando no sea buscada para su uso en un procedimiento penal de carácter potencial contra las personas que son realmente vigiladas. Esto tiene implicaciones para la aplicación de la Cuarta Enmienda en las investigaciones acerca de los terroristas. La Corte Suprema ha afirmado que "la Cuarta Enmienda casi con toda certeza permitiría un bloqueo al camino que se ajustara precisamente a la situación enfrentada y obedeciera al propósito de impedir un ataque terrorista inminente".[4] Esto se dijo sobre un caso en el que la Corte determinó que un bloqueo al camino *no* se podía permitir si solamente respondía al objetivo de tratar de atrapar a un traficante de drogas. Así pues, la Corte estaba reconociendo que una *L* más elevada puede justificar una *P* más baja; entre más severo sea el delito, menor será la causa probable que la policía pudiera requerir a fin de justificar un cateo o búsqueda de un determinado grado de invasión (*B*). Cuando en julio de 2005 fue bombardeado el sistema de tránsito de Londres, con un saldo de más de 50 personas muertas, la policía de la ciudad de Nueva York comenzó a efectuar cateos aleatorios de bolsas y paquetes en el metro a pesar de que, según las apariencias, no había pruebas de que se hubiera planeado o fuera inminente un ataque. El peligro de un ataque era ligero, pero también lo era el costo impuesto por unos cateos extremadamente leves. (Sobre la necesidad de modificar los procedimientos penales para combatir el terrorismo internacional, véase también antes VII.12.)

[4] City of Indianapolis *vs.* Edmond, 531 U.S., 32, 44 (2000).

XXX.2. Los remedios de la Cuarta Enmienda

Hay tres remedios comunes contra las violaciones de la Cuarta Enmienda. Uno es preventivo: el requerimiento de que, cuando sea viable, la policía obtenga una orden de detención emitida por un juez o un magistrado. En realidad, la Cuarta Enmienda no requiere con tantas palabras una orden de detención en ningún caso. La única referencia a las órdenes de detención es una prohibición de las órdenes de detención generales. Cuando la enmienda se aprobó, los funcionarios de aduanas y otros agentes gubernamentales solían tratar de obtener órdenes de cateo con la esperanza de que ello inmunizara al agente contra un juicio de allanamiento levantado por la persona cuya propiedad habría de catearse o confiscarse. Apenas en el decenio de 1940 la Corte Suprema decidió que la orden autorizada era una protección para el objetivo del cateo. La importancia de la orden autorizada consiste en que la determinación de la causa probable la hace o, más realistamente, la revisa el magistrado, que no es parte de la policía y, por ende, presumiblemente examina con más imparcialidad las pruebas que han inducido a la policía a decidir el cateo.

La Corte ha dado un giro de 180 grados al sostener que si los policías que ejecutan la autorización confían de buena fe en la aparente validez de la orden autorizada, serán inmunes contra la responsabilidad de los daños causados a la víctima si la orden fuese inválida por alguna razón. El problema de este enfoque es que la determinación de la causa probable la hace un magistrado en un procedimiento aparte, antes que un juez y un jurado en un juicio de daños, como se hace cuando no hay orden autorizada. El remedio de los daños se aproxima más a un enfoque de mercado; la determinación del magistrado se aproxima más a un enfoque burocrático, al esfuerzo para impedir los cateos ilegales.

El remedio más importante contra los cateos ilegales es todavía la regla excluyente, en cuyos términos las pruebas y los indicios conducentes a pruebas producidos por un cateo ilegal no podrán usarse como pruebas en un juicio penal contra la persona cuyos derechos fueron violados. Esta regla es controvertible pues puede generar una disuasión excesiva de las investigaciones penales; y, en efecto, hay evidencias empíricas de que la imposición que hizo la Corte Suprema de la regla excluyente en los estados suscitó un incremento en la tasa de delitos.[5] Supongamos que B, el costo para el acusado del cateo en términos de los daños causados a la propiedad o de la intervención de comunicaciones privadas legítimas, es de 1 000 dólares; P, la probabilidad de que no pudiera ser condenado sin este cateo, era de 1% en el momento del cateo; y L, el costo social (de una menor disuasión y prevención del delito) de la

[5] Véase Raymond A. Atkins y Paul Rubin, "Effects of Criminal Procedure on Crime Rates: Mapping Out the Consequences of the Exclusionary Rule", 46 *J. Law & Econ.*, 157 (2003).

ausencia de condena, es de 50 000. Por lo tanto, el cateo violará la Fórmula de Hand. Pero supongamos que las pruebas obtenidas en el cateo son esenciales para la condena. Esto no es incongruente con el hecho de que *P* fuera muy baja en el momento del cateo. Puede haber sido baja porque la policía no tenía buenas razones para pensar que el cateo sería productivo —era un palo de ciego— y no porque hubiera métodos alternativos, menos invasores de la privacidad, para obtener pruebas esenciales. Por lo tanto, aunque el costo social del cateo es de sólo 1 000 dólares, la regla de exclusión impondrá un costo de castigo de 50 000 a la sociedad. Esto omite el costo para el acusado de ser castigado, pero tal omisión es correcta porque habrá sido incluido en la estimación de que el costo social de la ausencia de castigo sería de 50 000 dólares.

Los costos sociales frecuentemente excesivos de la regla excluyente se reconocen en algunos casos. Por ejemplo, si los policías arrestan ilegalmente a un hombre, ello no les impide someterlo a juicio, aunque sí les está vedado el uso de cualquier prueba obtenida de él a resultas del arresto. El costo de no ser capaces de someterlo a juicio no superaría el costo de tener que renunciar al empleo de algunas pruebas, aunque los dos costos convergen cuando la prueba es esencial para la condena.

Por supuesto, la hipótesis de la disuasión excesiva supone que hay otros remedios para los cateos ilegales, además de la regla excluyente. Si la regla excluyente fuese el único remedio y fuera abolida, habría demasiados cateos ilegales porque los costos para las víctimas no entrarían en el cálculo de la decisión de la policía y los fiscales. ¿O los habría? Si la regla excluyente se traduce en una tasa de delitos más elevada, la policía practicará más cateos, algunos (tal vez muchos) de los cuales serán ilegales.[6]

Sin embargo, hay un remedio alternativo: el juicio de cuasidelito por un cateo o una confiscación inconstitucionales. La sección 1 de la Ley de Derechos Civiles de 1871[7] autoriza los juicios de cuasidelito para hacer valer los derechos constitucionales federales contra los funcionarios estatales, y juicios similares pueden promoverse directamente de acuerdo con la Cuarta Enmienda contra funcionarios federales. Un juicio de cuasidelito es preferible a la regla excluyente desde un punto de vista económico porque permite que la sanción sea proporcional a los costos sociales efectivos de la invasión de la privacidad causada por el cateo ilegal. En efecto, la policía tendría que "comprar" a la víctima los frutos de sus cateos "ilegales", lo que haría presumiblemente cuando los beneficios probatorios excedieran los costos del cateo para la víctima. Por lo tanto, se excluirían menos pruebas que con la regla excluyente, que deja fuera las pruebas ilegalmente confiscadas aun si los beneficios alcanzados en la aplicación de la ley de un cateo ilegal en particular exceden a los costos que enfrenta el demandante respecto a daño de propiedad o pérdida de la privacidad.

[6] Hugo Mialon y Sue H. Mialon, "The Effects of the Fourth Amendment: An Economic Analysis", 24 *J. Law, Econ. & Organization,* 22 (2007).

[7] 42 U.S.C., 1983.

Con todo, hay una preocupación de que el remedio del cuasidelito pueda ser ilusorio en muchos casos. Los jurados sienten a menudo escasa simpatía hacia los delincuentes, de modo que podrían sentirse renuentes a otorgar daños sufridos por un delincuente en el curso de un cateo ilegal que sin embargo rindiera pruebas positivas de sus actividades delictivas. Una dificultad mayor aún es el problema de las inmunidades frente a la responsabilidad por daños. La regla solía ser que el demandado (en el juicio de cuasidelito, es decir, el funcionario que aplica la ley) era inmune si había creído razonable pero incorrectamente que el cateo era legal. Según los principios ordinarios del cuasidelito, esto no sería una defensa, pues la responsabilidad tiene el efecto saludable de otorgar a las personas un incentivo para descubrir qué dice la ley. Puede sostenerse que, dado que es improbable que los policías capten el beneficio social pleno de un trabajo meticuloso, tampoco debieran pagar los costos sociales plenos de una meticulosidad excesiva. Sin embargo, este argumento es débil, porque, incluso si los policías fueran plenamente responsables, su empleador podría restablecer los incentivos previos aceptando indemnizarlos por la responsabilidad cuasidelictiva, como en efecto se hace de ordinario. La ley actual es ligeramente más favorable para las víctimas de los cateos ilegales. Para ser inmune, el policía debe haber actuado con buena fe "objetiva", la cual se define como la conformidad con la ley claramente establecida, pero sus creencias subjetivas son irrelevantes. Esto protege contra la imposición de la responsabilidad por daños con base en nuevas teorías legales de la ilegalidad.

¿Debiera imputarse la inmunidad del policía al empleador —el departamento de policía estatal o local— si es el demandado? Lógicamente no, pues el departamento está en una posición mucho mejor que el policía individual para internalizar los beneficios de un trabajo policiaco vigoroso. Pero la ley existente no permite que el demandante considere responsable al empleador en un juicio de derechos civiles sobre la base de que el superior responde. Eso tiene sentido económico; cualesquiera argumentos que estén disponibles para el *respondeat superior* en los casos de cuasidelitos (véase antes VI.8) se aplican igualmente a los casos de cuasidelitos constitucionales.

Si el remedio de cuasidelitos para un cateo inconstitucional es efectivo, la regla excluyente no puede ser justificada desde una perspectiva de costo-beneficio. La razón de ello es que subsidia la actividad delictiva. Los únicos beneficiarios de la regla (dando por supuesto que el remedio de cuasidelitos alternativo tenga eficacia) son los delincuentes, dado que la regla excluye solamente el contrabando u otra evidencia de delitos. La condición de que el remedio de cuasidelitos sea una alternativa adecuada es, no obstante, crucial, pues en su ausencia haría falta la regla excluyente para proteger a las personas inocentes de cateos policiacos que fueran irrazonables, toda vez que la policía no pagaría ningún precio si el cateo no lograra hallar material de contrabando u otra prueba de delito.

XXX.3. El enigma de la autoincriminación y de confesiones por coerción

La Quinta Enmienda crea un derecho de no ser obligado a autoincriminarse. La importancia principal del derecho es como un privilegio probatorio (véase antes XXIII.8). Una persona no puede ser obligada a testificar si su testimonio pudiera demostrar que ella había cometido un delito. Tampoco puede un fiscal pedir al jurado que infiera la culpa del acusado a partir de su ejercicio del privilegio de la Quinta Enmienda. Tampoco se permite que el gobierno utilice una confesión que han extraído de una persona mediante la tortura u otros medios ilegales como prueba en contra suya.

Una cuestión inicial es la eficacia que tiene el privilegio al otorgar a una persona una elección libre acerca de prestar o no un testimonio autoincriminatorio. La teoría de juegos debiera alertarnos acerca de un problema (del tipo que exige un desentrañamiento) generado por el derecho a renunciar a un privilegio probatorio. Supongamos que sólo los culpables harían la clase de admisiones perjudiciales que el privilegio contra la autoincriminación forzosa (como los privilegios abogado-cliente y los maritales; véase antes XXIII.8) permite que los acusados eviten u oculten. Entonces el inocente renunciaría siempre al privilegio a fin de señalar su inocencia.[8] Quienquiera que no lo hiciera sería considerado culpable con razón. Ésta es la base para que los jurados infieran a menudo la culpa a partir de la negación del acusado a subir al estrado, aunque se les instruya para que no obtengan ninguna inferencia de culpa a partir de su negativa. Con todo, si el inocente soporta también un costo al renunciar al privilegio, la negativa a la renuncia no puede interpretarse confiablemente como una señal de culpa. Los jueces que desean que los jurados se tomen en serio el principio de que la culpa no debiera inferirse de una negativa a renunciar al privilegio de un demandado de no ser obligado a autoincriminarse tendrían que ofrecer a los jurados una explicación creíble de la razón por la que una persona inocente podría temer las consecuencias de testificar.

La decisión del acusado penal acerca de testificar, renunciando a su privilegio contra la autoincriminación forzosa, puede modelarse como $p = p_1x_1 + (1 - t)p_2 + tp_3tx_2$, donde p es la probabilidad de que el acusado sea declarado culpable, p_1 es la probabilidad de la culpa generada por las otras pruebas del caso (x_1), p_2 es la probabilidad de la culpa que el jurado inferirá si el acusado no sube al estrado, p_3 es la probabilidad de la culpa que inferirá el jurado si el acusado sube al estrado y testifica (siendo x_2 el testimonio que prestará), y siendo t la decisión de testificar o no, tomando t un valor de 1 si testifica y de 0 si no lo hace. Si testifica, el término medio del miembro derecho de la

[8] Véase en general Daniel R. Fischel, "Lawyers and Confidentiality", 65 *U. Chi. L. Rev.,* 1 (1998).

ecuación desaparece, pero si su testimonio es perjudicial para él entonces el tercer término (que se vuelve simplemente p_3x_2) será positivo, mientras que si se niega a testificar desaparece el tercer término pero el segundo término es positivo. Por lo tanto, la decisión de testificar dependerá de una comparación del término medio en el caso de la negativa con el tercer término en caso de la aceptación de testificar.[9]

No es fácil encontrar una justificación económica de la cláusula de autoincriminación. Hobbes pensaba que, puesto que los hombres habían formado el contrato social constitutivo del Estado a fin de salvar sus vidas (vidas que en el estado de naturaleza eran, según Hobbes, solitarias, pobres, sucias, brutales y breves), tenían derecho a rescindir el contrato si el Estado amenazaba sus vidas, de modo que no tenían ninguna obligación de acusarse a sí mismos, por lo menos de un delito capital. Éste es un mal argumento desde un punto de vista económico. Si una regla de autoincriminación forzosa, al volver más eficiente el cumplimiento del derecho penal, reduce la tasa de homicidios, podríamos suscribir *ex ante* la regla, aunque *ex post* nos viéramos perjudicados por ella porque resultáramos ser homicidas antes que una víctima potencial del homicidio.

Un argumento mejor para explicar el privilegio es que otorga al acusado culpable una elección entre subir al estrado y mentir, por una parte, y guardar silencio, por la otra. Si se eliminara la segunda opción aboliendo el privilegio, los acusados culpables subirían al estrado con mayor frecuencia y mentirían más a menudo. Dado que la mayoría de los acusados penales son culpables, la mayoría de los acusados penales que testifiquen mentirán, y esto volvería más difícil que a los acusados inocentes se les creyera; los investigadores de los hechos esperarían que el testimonio de los acusados fuese deshonesto.[10]

Aun si no hubiera ninguna prohibición absoluta de que se obligue a las personas a autoincriminarse —de modo que, por ejemplo, un fiscal estuviera en libertad de comentarle al jurado sobre la renuencia del acusado a subir al estrado, o que el acusado pudiera ser condenado por desacato por negarse a testificar y contestar todas las preguntas que se le formulen—, todavía habría razones económicas para no permitir (en general) el empleo de la tortura para obtener confesiones. Podemos comparar los beneficios de una confesión al reducir el costo de la condena de una persona culpable con el costo de ex-

[9] Por cierto, este enfoque puede emplearse para modelar cualquier caso en el que la ausencia de pruebas genere una inferencia, como en el caso hipotético del autobús discutido antes en XXIII.4. En ese caso, p sería la probabilidad de que el propietario del autobús que lesionó al demandante fuese la Compañía A, el demandado; p_2 sería la contribución (negativa) a p si la única prueba es el porcentaje de los autobuses propiedad de A y si esto implica que el demandante está reteniendo pruebas de que el autobús en cuestión era en realidad propiedad de B (es decir, si $t = 0$); y p_3 sería la contribución a p si el demandante aporta pruebas adicionales (x_2), fuera de la mera estadística, acerca de la propiedad.

[10] Daniel J. Seidmann y Alex Stein, "The Right to Silence Helps the Innocent: A Game-Theoretic Analysis of the Fifth Amendment Privilege", 114 *Harv. L. Rev.,* 430 (2000).

traer la confesión, el costo de aumentar la probabilidad de condenar a una persona inocente, y el costo del dolor, el sufrimiento y la humillación de la persona torturada. Cuando analizamos el cateo y la confiscación vimos que la justificación de los métodos policiacos costosos es una función positiva de la gravedad del delito que se investiga, pero esta relación se oscurece en el contexto de la confesión por el hecho de que la condena de la persona equivocada es más costosa entre más grave sea el delito. La variable clave en este contexto es el grado de la brutalidad empleada contra el sospechoso, porque influye sobre los tres costos, y en la misma dirección. Entre mayor sea la brutalidad, más probable será que una persona inocente confiese, mayor será el sufrimiento experimentado por la persona interrogada y mayor será el costo del interrogatorio para los interrogadores a menos que sean sádicos.

No obstante, esto da por sentado, como lo hizo el análisis del cateo y la confiscación en la primera sección de este capítulo, que el objetivo del interrogatorio (o del cateo) es obtener evidencias para una acusación penal. Esto no es siempre cierto. El propósito podría ser sencillamente impedir que sobrevenga un delito, o incluso solamente descubrir que pende una amenaza. Tal vez no se le aplique ningún castigo a la persona objeto del cateo. De hecho, es más probable que los juicios de daños por violación de los derechos de la Cuarta o la Quinta Enmienda se presenten ante el tribunal, y que tengan éxito, si, por la razón que sea, el demandante no es perseguido. Los costos de los cateos y los interrogatorios deben ser refigurados cuando la acusación penal no ocupa el proscenio, o lo abandona. El análisis costo-beneficio del interrogatorio coercitivo se vería alterado dramáticamente si, por ejemplo, el interrogatorio concerniera a un complot terrorista y la persona interrogada —una figura periférica del complot pero, aun así, poseedora de información vital— no se enfrentara a ningún castigo penal sino únicamente a la deportación en calidad de inmigrante ilegal, a la vigilancia continua o a una advertencia.

Además, la economía del interrogatorio implica que los tribunales debieran ser más tolerantes con el empleo de las declaraciones falsas y otros trucos para extraer confesiones de lo que son con la brutalidad. Una promesa falsa ("Confiesa y no se te castigará") podría inducir a una confesión falsa, pero no impone costos sustanciales al interrogado ni a los interrogadores, de modo que es más probable que se justifique por el costo que la brutalidad física. O veamos el caso común donde, para inducir una confesión, la policía exagera las otras pruebas que tiene de la culpa del sospechoso. Con esta exageración está tratando de persuadirlo de que el costo de la confesión para él es menor de lo que es en realidad. Es improbable que esta táctica provoque una confesión falsa; y, en un caso donde la policía no ha podido obtener otras pruebas de la culpa, el beneficio de la confesión se maximiza, mientras que si la policía tiene muchas otras pruebas no necesitará exagerarlas. Por lo tanto, el análisis costo-beneficio apoya fuertemente que se permita a la policía y a los fiscales el empleo de esta táctica, y los tribunales realmente lo permiten. Pero en total

restringen el interrogatorio más de lo que parece justificado por el costo. Así que el derecho y la economía no son enteramente congruentes después de todo. Pero eso ya lo sabía usted.

Lecturas recomendadas

Campbell, Charles F., Jr., "An Economic View of Developments in the Harmless Error and Exclusionary Rules", 42 *Baylor L. Rev.,* 499 (1990).
Fischel, Daniel R., "Lawyers and Confidentiality", 65 *U. Chi. L. Rev.,* 1 (1998).
Posner, Richard A., *Not a Suicide Pact: The U.S. Constitution in a Time of National Emergency*, cap. 4 (2006).
——— "Rethinking the Fourth Amendment", 1981 *S. Ct. Rev.,* 49.
——— "Excessive Sanctions for Governmental Misconduct in Criminal Cases", 57 *Wash. L. Rev.,* 635 (1982).
Seidmann, Daniel J. y Alex Stein, "The Right to Silence Helps the Innocent: A Game-Theoretic Analysis of the Fifth Amendment Privilege", 114 *Harv. L. Rev.,* 430 (2000).

Problemas

1. Una persona *racional,* ¿confesaría alguna vez *voluntariamente* haber cometido un delito?

2. ¿Es congruente la discusión en este capítulo del remedio del cuasidelito para las violaciones a la Constitución con la discusión en el capítulo III de la compensación justa por las confiscaciones?

3. ¿Cuáles serían los costos y beneficios sociales de modificar la regla excluyente a fin de requerir la exclusión de las pruebas obtenidas ilegalmente sólo si la policía que llevara a cabo la búsqueda cuyo fruto fue la evidencia hubiera sabido que dicha pesquisa y la reunión de evidencias resultante violaban la Cuarta Enmienda?

Octava Parte

DERECHO COMPARADO E INTERNACIONAL

XXXI. DERECHO COMPARADO E INTERNACIONAL

XXXI.1. Derecho comparado

El concepto "derecho comparado" por lo general se utiliza para referirse al derecho interno de las naciones extranjeras, enfatizando las diferencias entre el derecho extranjero y el estadunidense. Ya se trataron las diferencias sistémicas más importantes en nuestro análisis del derecho y la modernización (véase antes VIII.6), del contraste entre el sistema de código del derecho civil y el derecho común angloestadunidense (II.3), del contraste entre el aspecto inquisitorial de la adjudicación del derecho civil y el aspecto acusatorio de la adjudicación estadunidense (XX.1, XIII.2), y, más ampliamente, del contraste entre los poderes judiciales de Europa Continental y los Estados Unidos (XX.1). En un espectro aún más amplio, los sistemas jurídicos mundiales se pueden dividir en unas cuantas familias, la más grande de las cuales es la que se conforma por los países angloparlantes sistemas jurídicos que provienen de la Inglaterra medieval; y los sistemas jurídicos (que incluyen aquellos de algunos países asiáticos, así como europeos y latinoamericanos) que provienen de sistemas franceses y alemanes, y, en última instancia, del derecho romano y canónico.[1]

Hay muchas diferencias sustanciales entre las "familias", pero sólo mencionaré aquellas relacionadas con el derecho contractual y el derecho antimonopólico. El enfoque de Europa Continental a los contratos se construye alrededor de la consigna *pacta sunt servanda* (lo pactado obliga), que es lo opuesto a la teoría de la opción de las obligaciones contractuales de Holmes que se abordó en el capítulo IV; según la teoría de Holmes, no existe la obligación de cumplir un compromiso contractual, sólo hay una obligación de pagar daños a la otra parte en caso de que se decida no ejecutarlo. La diferencia en el enfoque tiene consecuencias prácticas en la disposición de los tribunales en el continente a ordenar el ejercicio específico de los compromisos contractuales. Esto es extraño en nuestro sistema. Y también es ineficiente. La ruptura de un contrato maximiza su valor si la parte que rompe el acuerdo puede com-

[1] Véanse Rafael La Porta, Florencio Lopez-de-Silanes y Andrei Shleifer, "The Economic Consequences of Legal Origins", 46 *J. Econ. Lit.,* 285 (2008); Nuno Garoupa y Mariana Pargendler, "A Law and Economics Perspective on Legal Families" (2013), próximo a publicarse en *The Methodologies of Law and Economics* (T. S. Ulen [comp.]); Michael L. Smith, "Deterrence and Origin of Legal System: Evidence from 1950-1999", 7 *Am. Law & Econ. Rev.,* 350 (2005); Klaus Gugler, Dennis C. Mueller y B. Burcin Yurtoglu, "Corporate Governance and the Returns on Investment", 47 *J. Law & Econ*., 589 (2004).

pensar por completo a la otra, ya que eso implica que el rompimiento (si es voluntario y no obligado por las circunstancias) debe de haber superado, en el sentido de Pareto, al cumplimiento del contrato. Pues esto se debe a que al menos una persona obtuvo ventaja de ello: la parte que rompió el contrato, y probablemente el tercero con el que estableció un nuevo contrato, y nadie, ni siquiera la parte con quien se rompió el acuerdo (si al final se le compensó por completo por ello), quedó en una situación peor. Además, como se mencionó en el capítulo IV, los tribunales de Europa Continental obligan a incluir cláusulas de penalización en los contratos; las cortes estadunidenses no. Una cláusula de penalización, como ejecución específica, puede evitar un rompimiento eficiente.

De esta forma, parece que el derecho contractual en Europa Continental ha sido en efecto influenciado por preocupaciones morales que son ajenas al derecho estadunidense y que quizá se originan en el derecho canónico. Se puede encontrar evidencia adicional en la noción continental más amplia del deber de buena fe en la ejecución de contratos,[2] y en el énfasis que el derecho continental pone en la culpa, algo que no se encuentra en el derecho común de tipo contractual. Esto pone a una parte en rompimiento de su contrato, y, por tanto, acreedora a una sanción legal sólo si se considera que "se podía esperar, con razón, que actuara de otra manera", es decir, sólo si él tiene la culpa al no cumplir el trato.[3] Por tanto, el derecho francés está más interesado en definir las intenciones de ambas partes que el derecho estadunidense. Así, si una de las partes tenía la intención de hacer una promesa, pero por alguna razón no logró plasmarla por escrito de forma inteligible, no se le permitirá romperla refiriéndose a un escrito. Hay una mayor "disposición [de las cortes francesas] para recurrir a negociaciones previas y acciones subsecuentes de los involucrados" al interpretar un contrato.[4]

Podemos ver diferencias más evidentes entre las leyes antimonopólicas de los Estados Unidos y la Unión Europea. En el capítulo X vimos que un juez estadunidense tiene dificultad para desprenderse de la idea de que deben existir acuerdos explícitos entre competidores como una condición previa para dar un veredicto de colusiones ilegales. Las leyes antimonopólicas de la Unión Europea (a la que se les llama derecho competitivo) utilizan el concepto de "dominio" para prohibir conductas de colusión tácita.[5] Las empresas con grandes acciones en el mercado de valores, aunque se quedan cortas de ser

[2] Ejan Mackaay, "The Civil Law of Contract", en *Contract Law and Economics,* 424 (Gerrit De Geest [comp.], 2011).

[3] Jürgen Basedow, "Towards a Universal Doctrine of Breach of Contract: The Impact of CISG", 25 *Int'l Rev. Law & Econ.,* 487, 495 (2005) ("el principio de culpa suele considerarse parte indispensable de la ley de obligaciones en países con derecho civil").

[4] Stefan Vogenauer, "Interpretation of Contracts: Concluding Comparative Observations", en *Contract Terms,* 123, 150 (Andrew Burrows y Edwin Peel [comps.], 2007).

[5] Marilena Filippelli, *Collective Dominance and Collusion: Parallelism in EU and US Competition Law*, 284-290 (2013).

monopolios, son sospechosas de colusiones tácitas, y se las puede obligar a reducir sus acciones sin haber hallado una colusión explícita, aun cuando también existe la sospecha bien fundamentada de que un veredicto de "dominio" se emite para proteger a las firmas pequeñas contra competidores que son grandes por su eficiencia. Ése es un elemento proteccionista en las leyes antimonopólicas europeas.

Aunque todavía persisten muchas de las diferencias procedimentales y sustanciales entre el derecho estadunidense y el de Europa Continental, existe una tendencia reciente de ambos a convergir, en particular en el procedimiento. Los jueces estadunidenses ahora necesitan alegatos más elaborados que la presentación de la demanda simplificada que se encuentra en las Reglas Federales del Procedimiento Civil (FRCP), y han cobrado un carácter más directivo y menos arbitral, y en todos esos aspectos son más similares a sus contrapartes extranjeras. Algunos jueces estadunidenses incluso han estado designando testigos periciales (véase antes XXIII.6), una inclinación al enfoque "inquisitorial" de los jueces en el continente, y al que se oponen muchos jueces estadunidenses al considerarlo incongruente con nuestra tradición "acusatoria". Por su parte, los jueces continentales han comenzado a ejercer mayor discrecionalidad que antes y han ganado algo del prestigio de sus similares estadunidenses, cuando tradicionalmente se había considerado a los poderes judiciales del continente casi indistinguibles de una función pública no judicial.[6]

La tendencia a convergir se debe en gran medida a la globalización creciente, y en particular al aumento de despachos jurídicos internacionales. La globalización tiende a promover la uniformidad internacional; por ejemplo, las orquestas de hoy en día suelen carecer de un distintivo nacional, ya que los integrantes de las grandes orquestas compran los mejores (y por lo general los mismos) instrumentos en donde sea que se produzcan en el mundo. En el ámbito legal, un efecto de la globalización es el aumento en el costo de patentes porque los vendedores del mismo producto patentado en diferentes países tendrán que obtener la patente en cada país en el que vendan si quieren retener el derecho a excluir a la competencia, lo que se traduce en costos legales considerables, acompañados de una mayor incertidumbre.[7]

Hay que considerar una peculiaridad de los procedimientos estadunidenses al tratar con el derecho extranjero. Cuando las leyes que se aplican para el caso de un tribunal estadunidense son extranjeras, incluso si se trata de un país en el que el idioma oficial es el inglés y cuyo sistema jurídico tiene el mismo origen, es decir, el derecho común inglés, nuestros tribunales por lo general basan su conocimiento del derecho extranjero en el testimonio de los

[6] Véase Scott Dodson y James M. Klebba, "Global Civil Procedure Trends in the Twenty-First Century", 34 *Boston College Comp. & Int'l L. Rev.*, 1 (2011).

[7] Véase, por ejemplo, Malwina Mejer y Bruno van Pottelsberghe de la Potterie, "Economic Incongruities in the European Patent System", 34 *European J. Law & Econ.*, 215 (2012).

abogados en lugar de tratados o artículos de revistas jurídicas. A los abogados que testifican acerca de cómo interpretar la ley extranjera, sean practicantes o profesores, se les paga por el testimonio y se les elige con base en la convergencia de su visión con la posición litigante del cliente o su disposición de ajustarse a la visión que los impele a adoptar el cliente. Esto pone en jaque al testimonio pericial. Cuando el testimonio se relaciona con un problema científico u otro asunto técnico, podría considerarse irracional esperar que un juez resuelva el asunto sin la ayuda de este testimonio. Sin embargo, los jueces son expertos en derecho, y hay material publicado en abundancia en forma de tratados, artículos en revistas jurídicas, estatutos y casos que arrojan luz sobre problemas del derecho extranjero en forma neutral.

XXXI.2. Los tratados y otros acuerdos internacionales

El resto de este capítulo se centrará en el derecho internacional, el derecho aplicable a relaciones y transacciones internacionales, tanto políticas como comerciales. El campo es vasto, mi alcance ligero, y los ejemplos a discutir pocos.[8]

Comencemos con los tratados, convenios internacionales (como los Convenios de Ginebra) y otros acuerdos entre dos o varias naciones, que por lo general no se analizan en términos de contratos, pero que sí *son* contratos. La economía de los contratos se puede aplicar a ellos con resultados fructíferos, aun si esto significa suponer que las naciones, al relacionarse unas con otras, ya sea de forma comercial o no, o incluso de modo beligerante, se comportan en gran medida como individuos en sus relaciones comerciales. Éste no es un supuesto tan descabellado como parece. La economía progresó mucho en la creación de modelos de interacción entre empresas de negocios al no echar un vistazo dentro de ellas, y en cambio considerarlas como individuos. El mismo enfoque agregado puede funcionar en el estudio de la interacción de las naciones. Por supuesto, las naciones difieren notablemente unas de otras, pero ocurre lo mismo con los individuos. La función de utilidad de

[8] El derecho marítimo es una rama importante del derecho internacional del que he hecho mención reiteradamente en el libro, pero que no analizo de manera sistemática. Algunas características notables de él son su aspecto *in rem*, es decir, una demanda consecuencia de una colisión marítima u otro agravio contra la navegación puede realizarse en cualquier puerto en el mundo en donde se encuentre el navío para facilitar la jurisdicción; asimismo, cualquier nación puede procesar a un pirata, también para facilitar la jurisdicción. (Los lineamientos acerca de piratería han sido aumentados para abarcar las violaciones a los derechos humanos, lo que ahora se puede procesar en cualquier país que capture físicamente al infractor aun si ninguna de las víctimas está relacionada con ese país.) Otra característica del derecho marítimo que vale la pena mencionar es la distinción que se hace entre aguas territoriales e internacionales, lo que evita que las naciones impidan la navegación al designar divisiones en los océanos, que es diferente a establecer un límite de área marítima adyacente reservada para la pesca y la minería.

un individuo para el que el honor es el valor máximo, lo que lo hace propenso a retar a otros hombres a duelo, es diferente del almacenista para el que la prosperidad es el valor máximo; e igualmente una nación gobernada por la primera clase de hombres tendrá una función de utilidad diferente a la de una nación de almacenistas.

Las naciones se relacionan unas con otras no sólo como competidores, sino también como compradores y vendedores, es decir, partes contractuales, y este tipo de relaciones dobles también son comunes entre empresas de negocios. Además, así como en las empresas hay competidores que ocasionalmente se coluden, las naciones también llegan a formar alianzas. Una parte de las prácticas contractuales que se asocian en particular con el comercio internacional también son recurrentes en las relaciones de negocios comunes y corrientes. Un ejemplo de ello es la cláusula de la "nación más favorecida". Supongamos que un fabricante le promete a un distribuidor que no le cobrará una cantidad menor a otro distribuidor. Esta promesa alienta al distribuidor a enfocarse en distribuir los productos de dicho fabricante, pues no se preocupará de que lo socaven otros distribuidores, aunque puede[9] que el distribuidor tenga que compensar al fabricante por comprometerse con él de esta manera, por ejemplo, al pagar un precio de venta al mayoreo más elevado del que estaría dispuesto a pagar sin la cláusula de la nación más favorecida. De la misma forma, una nación importadora que otorga un estatus arancelario de tipo nación más favorecida a una exportadora podría, al hacerlo, alentar al exportador a que se enfoque en la producción para el mercado de importación, pero también incurrirá en mayores costos en caso de que tiempo después quisiera otorgar una rebaja arancelaria a otra nación exportadora.

Las relaciones internacionales suelen pensarse como diferentes en sumo grado de aquellas entre individuos y empresas ubicados en una nación tan sólo por la carencia de una autoridad supranacional fuerte que haga cumplir las leyes y normas internacionales. Pero hemos visto que, aun en el círculo nacional, gran parte del cumplimiento de contratos se encuentra descentralizado. La mayoría de los contratos se acatan, no por temor a acciones legales, sino por preocupaciones sobre la reputación y debido a la disponibilidad de remedios de autoayuda, como la recuperación y otros medios de embargo, así como la práctica de retener la ejecución recíproca (o los pagos). Los cárteles ilegales, como vimos en el capítulo X, ejemplifican acuerdos contractuales que no se pueden hacer cumplir de ningún modo por medio de una autoridad central; no obstante, pueden disfrutar de cierta estabilidad a consecuencia de un reconocimiento del interés mutuo y de amenazas de venganza que no ne-

[9] "Puede tener", en lugar de "tendrá", porque el productor puede buscar estar sujetado a una cláusula de "nación más favorecida" si sus competidores también están sujetos a una; de esta manera, al hacer que el recorte de precios sea más costoso, podría facilitarse el establecimiento en todo el mercado de tarifas superiores a las competitivas, como sabemos de antes, X.1. Esto se discute con más detalle en el capítulo acerca del antimonopolio.

cesitan ser explícitas. La venganza también desempeña un papel secundario importante en el cumplimiento de la ley penal, ya que por lo general las víctimas del delito no cuentan con el incentivo económico para cooperar con la policía y los fiscales a fin de ejercer la justicia contra los delincuentes.

Las convenciones internacionales sobre la guerra, que abarcan temas como el trato a los prisioneros, la ilegalización de ciertas armas y los derechos de las partes neutrales, abren preguntas interesantes. Una de ellas es por qué las reglas sobre la guerra suelen promulgarse en forma de convenciones internacionales, es decir, de manera multilateral en vez de mediante acuerdos bilaterales entre las naciones. Esto tiene relación con la distinción económica entre derechos contractuales y derechos de cuasidelitos. Un contrato te da derecho sólo sobre la otra parte contratante. Un derecho de los cuasidelitos, al igual que un derecho de propiedad (para cuyo cumplimiento a menudo sirven los derechos de cuasidelitos), es un derecho en contra de todo el mundo, pues otorga la posibilidad de cobrar daños, por ejemplo, a alguien que ha allanado una propiedad y con quien no se tiene un acuerdo previo que limite su derecho a entrar a ella. Sería impráctico elaborar un contrato con todo intruso potencial para protegerse de la entrada ilegal. De la misma forma, a no ser que una nación sepa con gran certeza quiénes son sus enemigos, ganaría muy poco de un acuerdo bilateral que limitara las conductas de guerra, pues éste no lo protegería de cualquier otra nación con la que podría encontrarse en guerra. Una convención internacional firmada por todos los beligerantes potenciales concede lo equivalente a un derecho de cuasidelitos, es decir, una protección contra todo el mundo (relevante). Aunque tiene un formato contractual, evita el costo de transacciones de varios contratos bilaterales al reunir a todas las partes potenciales bajo una sola negociación. A pesar de ello, los costos de transacción de una negociación entre múltiples partes son elevados, y por ello las convenciones sobre guerra suelen abarcar un solo tema, como los prisioneros o la seguridad de los civiles, en lugar de toda la gama de acciones a ejecutar durante el periodo de guerra que las naciones quisieran limitar recíprocamente.

Dada la incertidumbre con la que se desarrollan las guerras (una incertidumbre arraigada en el carácter intermitente y estratégico de las guerras, incluidos el valor estratégico de ocultar armas militares, tácticas y planes),[10] las convenciones de guerra suelen requerir una ejecución simétrica. El firmante de un convenio de este tipo no aceptaría abstenerse de maltratar a sus prisioneros a cambio de un pago complementario por parte de otro firmante, pues sería muy complicado calcular el valor monetario de la promesa del

[10] Hay que tomar en cuenta la implicación de que en ocasiones el espionaje puede resultar benéfico para la nación a la que se está espiando, o, en un sentido más generalizado, el mundo entero podría beneficiarse del espionaje. (*Cf.* el descubrimiento anterior al juicio, discutido en el capítulo XXII.) La regla en el derecho internacional que permite que un país nombre a un militar como parte de los integrantes de sus embajadas (es decir, un agregado militar) explícitamente permite cierto grado de espionaje.

primer firmante. Tampoco se intercambian *tipos* de limitación de armas: por ejemplo, una nación que acepte no embarcarse en la guerra submarina irrestricta a cambio de que otra no utilice gas venenoso. La ejecución puede variar cuantitativamente, como es el caso del número de acorazados que permitía el Tratado Naval de Washington de 1920, o el número de misiles nucleares permitidos por varios tratados entre los Estados Unidos y la URSS. Con todo, las consecuencias son más fáciles de estimar.

Los esfuerzos más efectivos a la hora de limitar guerras mediante convenciones internacionales parecen ser aquellos en los que los beneficios mutuos de la limitación son amplios (algo de consideración vital dada la dificultad de establecer pagos complementarios u otra forma de compensación en contratos multilaterales) y simétricos, como es el caso de convenciones que exigen el trato humanitario de los prisioneros. Los beneficios del trato humanitario a los soldados propios que han caído prisioneros son evidentes. Esto conlleva sus costos (los soldados propios podrían luchar con mayor fiereza si supieran que morirían en caso de rendirse); sin embargo, es complicado sacar ventaja de esta manera, pues el oponente puede responder de manera efectiva *no* maltratando a los prisioneros. Por lo tanto, salvo en contadas ocasiones, ninguna nación combatiente puede beneficiarse sustancialmente de maltratar a sus prisioneros. Si bien es cierto que la nación que mantiene más prisioneros incurre en gastos mayores para mantenerlos en condiciones favorables, este costo, que de cualquier manera no es tan grande, se ve compensado por el debilitamiento del enemigo que está implicado en el hecho de haber logrado capturar más prisioneros que el otro.

Este punto también explica por qué es más probable que las convenciones sobre el trato a prisioneros se hagan cumplir sólo por los signatarios aun si una nación combatiente no buscara mantener la reputación de cumplir promesas. Si ambas partes mantienen el mismo número de prisioneros, cada una de ellas tiene un medio simple y efectivo de tomar represalias en caso de que su enemigo maltrate a uno de ellos. Si una mantiene mucho menos prisioneros que la otra, lo más probable es que esté perdiendo la guerra y tema algún castigo si los maltrata; asimismo, tendrá menos beneficios, ya que el costo de mantener sólo a unos cuantos prisioneros será poco. El lado ganador, que tiene una cantidad desproporcionada de prisioneros, puede darse el lujo de mantenerlos, precisamente porque va ganando, y no obtiene mucho de maltratarlos, en especial si existe el riesgo, aunque quizá sólo sea pequeño, de que su oponente tome represalias con los prisioneros que tiene.

Aunque las convenciones acerca del trato humanitario de prisioneros generan beneficios mutuos, existe un costo social que suele omitirse con frecuencia: al reducir los costos de una guerra, las convenciones incrementan las posibilidades de una, tanto de forma directa como indirecta al facilitar el reclutamiento, lo que llevan a cabo reduciendo los costos esperados de convertirse en un combatiente. Comparemos la suposición generalizada de que, de

no ser por la posesión de armas nucleares por parte de la Unión Soviética y los Estados Unidos, las dos naciones habrían entablado una guerra entre 1945 y 1990. Esto implica que un tratado que proscribe dichas armas hubiera incrementado las posibilidades de una guerra, aunque ello también habría reducido los costos de una, por lo que el efecto del tratado en los costos esperados de guerra habría sido incierto.

El Tratado Naval de Washington representa un mejor ejemplo de una reducción de armas que beneficia a ambas partes (mejor, porque los barcos de guerra no tienen el mismo efecto disuasivo que las armas nucleares, por lo que una reducción de ellos tiene pocas posibilidades de afectar la disuasión). Los barcos de guerra son costosos, y si el enemigo puede construirlos tan rápido como usted, los gastos realizados no generarán ningún producto neto. El acuerdo generalizado de prohibir el uso de gases venenosos tras la primera Guerra Mundial ilustra otra situación en la que una limitante de guerra desemboca en un beneficio mutuo: el gas no resultaba un arma tan efectiva y podía, en todo caso, contrarrestarse de forma eficaz con máscaras de gas y otro tipo de vestimenta protectora, lo que indicaba que la guerra con gases era en no poca medida puros gastos y nada de beneficios. Esto es muy parecido al caso de los barcos de guerra, con la diferencia de que el costo de la "carrera armamentista" con gases era menor; empero, los beneficios percibidos también se consideraban menores. Se creía que eran tan pequeños que sólo bastaban los beneficios modestos de tener la reputación de mantener una promesa para inducir el cumplimiento de los convenios.

En donde uno espera que se rompa un convenio de control de armas es en una situación de guerra general, y no limitada. En una guerra limitada se espera que usted y su enemigo sobrevivan, y ello reduce los beneficios de ganar y los costos de perder, así y, por ende, los beneficios de obtener alguna ventaja por violar el convenio. En cambio, se espera que una guerra general lleve a la aniquilación del régimen de alguno de los combatientes, por lo que los beneficios que se obtienen de violar el convenio son mayores y, por tanto, se debe esperar que haya menor cumplimiento de éste.

La ausencia de autoridades supranacionales fuertes a cargo del cumplimiento de las leyes transforma a la reputación en un factor muy importante para el cumplimiento de acuerdos internacionales. Sin embargo, la reputación no es unitaria. Una nación puede tener buena reputación de cumplir con un tipo de compromisos y una mala con otro tipo.[11] No obstante, lo mismo es verdad para los individuos: aun si una persona se ha casado cinco veces, puede que sea muy escrupulosa al cumplir con sus obligaciones financieras. No hay necesidad de detectar irracionalidad en la variabilidad que muestran las personas o las naciones al honrar sus compromisos, y, en ese sentido, tampoco

[11] Véase George W. Downs y Michael A. Jones, "Reputation, Compliance and International Law", 31 *J. Legal Stud.*, S95 (2002).

hay necesidad de restarle valor a la hipótesis de la "caja negra" y considerar la estructura política interna de un país. La reputación es una predicción de lo que una persona o una nación harán en una situación determinada, y las violaciones anteriores de compromisos son sólo un tipo de datos que afectan a dicha predicción. Ya que los costos y beneficios por cumplimiento varían en consonancia con estos compromisos, violar uno de ellos podría representar evidencia débil, o aun nula, de que sea probable la violación de otro en donde los beneficios de cumplimiento podrían ser mucho mayores.

Con todo, de la misma manera, los beneficios de cumplir un compromiso específico pueden ser menores que los costos. En ese caso, el beneficiario de la promesa no cuenta con que el promitente tema perder su reputación y ello lo disuada de romper el acuerdo. Esto debería aumentar la demanda de otro tipo de sanciones en caso de una ruptura. A pesar de ello, las sanciones formales por rompimiento de un acuerdo internacional suelen ser bastante moderadas. Para entender el porqué, hay que considerar dos regímenes de sanción diferentes. Uno cuenta con castigos severos pero muchas defensas; el otro, con castigos moderados y ninguna defensa. Observaremos un contraste similar si comparamos las responsabilidades de un cuasidelito intencional y las responsabilidades por la ruptura de un contrato. Se pueden imponer sanciones punitivas a un cuasidelito intencional, pero ello requiere pruebas de que la conducta del acusado se podía evitar fácilmente: el significado operacional de infracción deliberada (véase antes VI.15). En efecto, la falta de malas intenciones funciona como defensa. La responsabilidad por el rompimiento de un contrato, no obstante, como ya se vio en este capítulo, es estricta; las defensas son contadas. Cuando la responsabilidad es estricta, las sanciones no deben ser punitivas, ya que provocan que se tomen cuidados excesivos para evitar la ruptura. Los acuerdos internacionales multilaterales involucran costos de transacción altos y una autoridad supranacional encargada de hacerlos cumplir, si existe, que por lo general será incapaz de tomar decisiones creíbles respecto a hechos controversiales. Así, esto conlleva a presionar a favor de la simplicidad y, por consecuencia, de pocas defensas, lo que se traduce en una responsabilidad por rompimiento más estricta que la relativa a la ruptura en contratos ordinarios, lo que implica que las sanciones óptimas ante cualquier infracción serán moderadas. Ya que se supone que estos acuerdos deben ser duraderos (pues de otra forma no valdría la pena incurrir en los altos costos de la transacción) y no se puede predecir el futuro, estas situaciones están llamadas a ser frecuentes, particularmente cuando el incumplimiento es inevitable e involucra costos económicos y políticos razonables… un motivo más para creer que las sanciones moderadas serían lo óptimo.

Un ejemplo relacionado es lo endeble de las sanciones por infringir los acuerdos realizados bajo el auspicio de la Organización Mundial del Comercio (OMC). El aspecto multilateral de estos acuerdos complica la renegociación y así abre paso a una argumentación a favor de sanciones débiles, cuidando de

evitar que un acuerdo imponga costos excesivos a un integrante. Ciertamente las sanciones son tan endebles que son casi imaginarias, en parte porque existen cláusulas de escape que permiten a las naciones integrantes, en condiciones particulares, suspender temporalmente sus obligaciones según el acuerdo sin ninguna sanción. Aun así, el cumplimiento razonable se consigue de todas formas, y es por razones políticas locales.[12] (Por lo que estamos abandonando la hipótesis de la "caja negra" por el momento.) Una disminución inicial de los aranceles, por medio del debilitamiento de la protección de la industria exportadora y el fortalecimiento de las industrias importadoras de insumos, mueve el poder político en favor de estas últimas, lo que, por consiguiente, hace más plausible la reducción subsecuente de los aranceles. En cuanto a las cláusulas de escape, una industria decadente se beneficiará más de un arancel que una industria floreciente porque, cuando la demanda va en caída, un aumento de precio no atraerá competidores nuevos. De esta manera, las cláusulas de escape tienden a beneficiar principalmente a industrias nacionales en declive a expensas de las competidoras extranjeras emergentes, lo que desembocará en una redistribución de la riqueza congruente con las presiones políticas.

El contrato bilateral privado ordinario tiene un periodo "natural", a saber, el tiempo anticipado de ejecución por la parte ejecutante. Este periodo se vuelve conflictivo en contratos multilaterales: entre más alto sea el costo de la negociación, más grande será la presión para establecer un periodo largo para el plazo del contrato, de modo que no se incurra en el costo con frecuencia. Además, los contratos multilaterales tienden, por su naturaleza, a crear relaciones permanentes y no *ad hoc*. (La Constitución de los Estados Unidos puede tomarse como un contrato de este tipo.) Entre más extenso sea el periodo, se debe crear mayor flexibilidad en la administración del contrato. Una forma de hacerlo es dejar los términos abiertos y hacer que un tribunal u órgano arbitral rellene los huecos de manera atenta a las circunstancias. De no ser esto factible, otra alternativa es permitir suficientes vías de escape y mantener sanciones por incumplimiento del acuerdo moderadas.

XXXI.3. EL *DUMPING* Y LA CUESTIÓN DEL LIBRE COMERCIO

El *dumping* hace referencia a una empresa extranjera que vende en los Estados Unidos a un precio menor del que cobraría en su mercado local. Esto está prohibido por la ley estadunidense (la Ley de Aranceles de 1930) cuando provoca "daños sustanciales" a la industria estadunidense. La Ley también impone "cuotas compensatorias" (impuestos) a los productos importados de los que se ha determinado que su producción se realizó con ayuda de un

[12] Warren F. Schwartz y Alan O. Sykes, "The Economic Structure of Renegotiation and Dispute Resolution in the World Trade Organization", 31 *J. Legal Stud.*, S179 (2002).

subsidio del gobierno del productor. Aunque normalmente no se considera a las regulaciones *antidumping* y a las cuotas compensatorias como leyes antimonopólicas, pretenden estar basadas en políticas similares. Y es cierto que la discriminación en costos y el subsidio gubernamental se alejan de precios competitivos y eficientes, aunque dicha conclusión depende del sentido específico con que se utilicen los términos "discriminación" y "subsidio".

La discriminación, en el sentido manejado por el economista, se refiere como sabemos al vendedor que cobra precios distintos sin justificarse en una diferencia en los costos; la competencia implica igualar el precio con los costos marginales, de modo que sólo una diferencia en el costo puede justificar una diferencia en el precio. Por lo general, una empresa puede discriminar sólo cuando tiene alguna influencia en el mercado, es decir, cuando se enfrenta a una curva de demanda de pendiente negativa, lo que es señal de que hay consumidores distintos dispuestos a pagar precios diferentes por el producto de la empresa. Y un subsidio diseñado para permitir que una firma venda por debajo del costo marginal genera el mismo tipo de distorsión que una influencia monopólica. El cambio de precio que permite este subsidio aleja a los consumidores de productos que son menos costosos de producir socialmente, tal como los precios monopólicos desvían a los consumidores hacia productos que son más costosos en su producción social.

Pero el hecho de que la discriminación y los subsidios puedan distorsionar la competencia no es una razón de peso para las leyes *antidumping* y de cuotas compensatorias. Supongamos que una empresa japonesa vende algún producto a precio monopólico en Japón y a precio competitivo, es decir, a un precio similar al costo marginal, en los Estados Unidos. Esto es discriminación de precios; empero, sólo puede afectar a las empresas estadunidenses si sus costos marginales son mayores que los de la empresa japonesa o si venden a precios monopólicos. En ninguno de los casos se beneficiarían la eficiencia, la competencia o el bienestar del consumidor si los Estados Unidos impusieran sanciones a la empresa japonesa. Si las industrias estadunidenses se ven afectadas cuando una empresa japonesa vende a un precio igual a su costo marginal, se trata de un daño autoinfligido, pues son las empresas estadunidenses las que no son capaces de reducir los costos o de competir.[13] El único argumento económico para sancionarla sería generar presión en el gobierno japonés para mitigar las limitantes competitivas en el mercado japonés que permiten que la empresa discrimine, pues estas limitantes podrían afectar a empresarios y trabajadores estadunidenses al evitar que las empresas estadunidenses entren al mercado japonés.

[13] De la misma manera, si los costos de una industria local se disparan, por lo que aumentan los precios y, en consecuencia, las importaciones aumentan y la industria local pierde ventas, ¿es el daño (la pérdida de ventas y, posiblemente, las ganancias) "culpa" de las importaciones o del incremento en los costos? Véase Donald J. Rousslang, "Import Injury in U.S. Trade Law: An Economic View", 8 *Int'l Rev. Law & Econ*., 117 (1988).

Por supuesto, si la empresa japonesa vende por debajo del costo en los Estados Unidos, con la esperanza de recuperar esos ingresos una vez haya erradicado a la competencia estadunidense, existe una preocupación antimonopólica normal. Pero cuando se trata de una industria extranjera completa, en lugar de una sola empresa, a la que se acusa de involucrarse en una fijación de precios depredadora en los Estados Unidos, dicha acusación deberá atenderse con aún más que el escepticismo ordinario que suscitan los cargos de precios depredadores. Una cosa es que una sola empresa decida reducir los precios primero y luego, quizá varios años después, ya que haya expulsado a sus competidores, aumentarlos. Que un grupo empresarial sea capaz de realizar semejante maniobra implica un grado de durabilidad y coordinación que pocos cárteles pueden lograr.

El análisis de las cuotas compensatorias es parecido. Si la empresa extranjera recibe un subsidio para vender en los Estados Unidos a un precio menor del que podría asignar normalmente, la pregunta económica es si el subsidio lleva el precio de la empresa por debajo del costo marginal, y, de ser así, si es probable que se retire el subsidio una vez que la competencia estadunidense de la empresa deje el mercado. De no ser el caso (si los exportadores de la empresa extranjera controlan las políticas económicas del país de tal modo que las exportaciones podrían recibir subsidios indefinidamente), existe una razón para tratar al subsidio como el otorgamiento de un beneficio permanente a los consumidores estadunidenses que es propenso a exceder los costos de los productores en los Estados Unidos. Los consumidores se benefician de un precio menor; los productores se ven afectados al inicio, pero presumiblemente pueden reubicar sus recursos (a largo plazo en todo caso) hacia otras industrias para beneficio de los consumidores. De otra forma se trata una vez más de una variante de la situación de los depredadores de precios.

Si un productor estadunidense comprueba que un productor extranjero vende por debajo del costo, ¿debería poder recuperar la pérdida total que se le impuso, o sólo la pérdida que habría sufrido si hubiese estado vendiendo a un precio igual al costo marginal y luego minimizara ese costo? Sería lo segundo, pues de otra forma el productor estadunidense se beneficiaría de haber cobrado a precios monopólicos o de haber fallado en minimizar sus costos.

Por supuesto, las preocupaciones que realmente animan a las leyes *antidumping*, las cuotas compensatorias y otras medidas contra las prácticas comerciales supuestamente "injustas" de productores extranjeros van bastante más allá de la preocupación por la depredación de precios. La preocupación principal es proteger a la industria estadunidense de los productores extranjeros con costos genuinamente menores, ya sea porque pagan salarios menores, incurren en menos controles de contaminación y otros costos reguladores, se administran mejor, tienen mejores empleados o disponen de plantas y equipos más modernos. ¿Existe un buen argumento económico que justifique las políticas proteccionistas? La mayoría de los economistas dirían que no, al menos

si se toma al bienestar económico *mundial* como el máximo; pues, en cambio, si el máximo es el bienestar de la nación, el proteccionismo es justificable en tres situaciones. Primero, si una nación es una gran importadora de un producto particular al grado de ser un monopsonio (véase antes X.12), podría, al imponer un arancel en el producto, obtener ingresos fiscales que, si se suman a los ingresos de los competidores locales protegidos, podrían superar los costos para los consumidores locales. No obstante, hay que destacar que una cuota, es decir, un tope en la cantidad de producto que puede importarse, no podría justificarse con esta base al no generar ingresos, aunque reduciría los precios si los Estados Unidos tuvieran un monopsonio del producto.

Aun si un país no cuenta con el poder para establecer un monopsonio, y aun si considera que el objetivo adecuado de una política es la maximización del libre comercio mundial, una medida proteccionista (mal llamada así en este caso) podría ser apropiada como una medida en represalia, para desalentar a otras naciones de adoptar medidas proteccionistas. Presumiblemente, entre más grande sea el poder de una nación en los mercados de importación mundiales, más poderosas serán sus amenazas de represalias. Por otro lado, un país así también podría tener capacidad de monopsonio, de modo que ya en la práctica sería complicado determinar si sus medidas proteccionistas en apariencia en verdad lo son o si se trata meramente de represalias. Los Estados Unidos, en sí un gran importador, ha explotado esta ambigüedad en sus políticas comerciales.

Por último, en caso de una depresión (o una recesión severa), el proteccionismo puede reiniciar la producción y el empleo, por lo menos a muy corto plazo. Reducir las importaciones, consecuencia de la promulgación de aranceles altos, aumenta la demanda de productos locales para remplazar las importaciones faltantes, y el aumento de la producción interna incrementa el número de empleos. No obstante, una medida más efectiva que no involucre aranceles u otras barreras explícitas al comercio sería devaluar la moneda nacional, con lo que las exportaciones se abaratan y las importaciones se encarecen. El incremento de la producción de artículos exportados aumenta la tasa de empleos, y la disminución de importaciones estimula la producción de sustitutos domésticos, por lo que el empleo también aumenta. Sin embargo, una devaluación en represalia por parte de los socios comerciales de ese país podría nulificar rápidamente los beneficios de la devaluación.

XXXI.4. La inversión extranjera directa

Como se acaba de analizar, el libre comercio puede ser aún motivo de controversia. Sin embargo, la inversión extranjera directa —comprar una empresa en un país extranjero o establecer una subsidiaria de una compañía en dicho país— provoca mucha más controversia, en especial si la empresa ex-

tranjera es china.[14] Un ejemplo de eso es el intento fallido ("frustrado" sería un término más apropiado) de parte de CNOOC, una compañía petrolera propiedad del gobierno chino, de comprar la empresa petrolera estadunidense Unocal en 2006. El intento provocó una tormenta política; se introdujeron leyes en el Congreso para prohibir la transacción, y CNOOC retiró la oferta.

El temor por la inversión extranjera directa proviene de un temor de que una nación extranjera le arrebate el control de industrias clave a un país (en este caso a los Estados Unidos), y que use su control para debilitar y dominar a la nación objetivo. Éste sería motivo de una legítima preocupación si a China u otra nación potencialmente hostil se le permitiera comprar en la industria de defensa estadunidense. No obstante, no lo es si pretende comprar una empresa petrolera u otra compañía que no produzca armamento militar.

China posee grandes cantidades de dólares estadunidenses. En la medida en que esas posesiones exceden las necesidades de mantener sus reservas financieras, la única manera de que China se beneficie de poseer dólares es si los intercambia por bienes de valor, tales como productos estadunidenses o activos de empresas estadunidenses. Y *queremos* que China realice esas transacciones, pues de otro modo tendrá menor disposición a retener dólares, es decir, menor disposición a vendernos bienes a cambio de pedazos de papel.

Se supone que una venta beneficia al vendedor y al comprador, por lo que deberíamos aceptar de buen grado la venta de compañías estadunidenses locales a compañías foráneas, de la misma manera que aceptamos la compra de bienes y servicios producidos en los Estados Unidos por parte de extranjeros. El miedo que alentó a gran parte de la oposición política a la venta propuesta de Unocal fue que el gobierno chino, de haber comprado la empresa, la hubiera obligado a vender todo su petróleo a China. ¿Pero por qué? Si lo hubiera vendido por debajo del precio, habría perdido dinero con la compra. ¿Y eso afectaría a los Estados Unidos? No mucho, y de ningún modo a largo plazo, pues el petróleo es una mercancía fungible, y simplemente tendría que remplazar el petróleo de Unocal con el de otros productores. No obstante, habría costos a corto plazo para recontratar el petróleo de remplazo. Pero una acción desafiante como ésta por parte de CNOOC sería improbable porque nosotros tomaríamos represalias. En este caso, el hecho de que casi una tercera parte de las reservas petroleras de Unocal se encuentran en América del Norte, antes que fortalecer el argumento a favor de bloquear la fusión, lo debilita, pues estas reservas funcionarían como un rehén en caso de que China utilizara su posesión de Unocal para perjudicar las reservas petroleras de los Estados Unidos.

El hecho de que China estuviera dispuesta a darnos un rehén semejante es una señal esperanzadora para el futuro de las relaciones entre China y Estados Unidos, pues sugiere que China prevé relaciones comerciales pacífi-

[14] Véase Angela Huyue Zhang, "Foreign Direct Investment from China: Sense and Sensibility" (King's College, Londres, 2013), http://ssrn.com/abstract=2307617.

cas y constructivas con Estados Unidos. De otro modo, no gastaría miles de millones de dólares para adquirir activos que en última instancia están bajo el control de nuestro gobierno.

XXXI.5. El Estatuto de Reclamación por Agravios contra Extranjeros

El Estatuto de Reclamación por Agravios contra Extranjeros[15] es una ley que delibera acerca de la jurisdicción de las cortes federales sobre "toda acción civil realizada por un extranjero que desemboque en un agravio o cuasidelito, y que atente en contra de la ley de naciones o algún tratado de los Estados Unidos". Un caso que resolvió mi corte[16] enfrentaba a 23 niños liberianos contra Firestone Natural Rubber Company, que operaba una plantación de caucho de 475 km^2 en Liberia a través de una subsidiaria y a la que los demandantes acusaban de violar el "derecho internacional consuetudinario" al emplear niños (aunque fuera indirectamente, pues imponía cuotas de producción que los adultos empleados no podían cumplir sin la ayuda de sus hijos) en el exigente y a veces peligroso trabajo de recolectar caucho.

La Corte Suprema ha dicho que "el derecho internacional es parte de nuestro sistema de derecho, y […] cuando no existe un tratado, y tampoco un acto ejecutivo o legislativo o una decisión judicial a cargo, debemos recurrir a los usos y costumbres de las naciones civilizadas".[17] El estatuto de reclamación de agravios o cuasidelitos contra extranjeros se promulgó en 1789, cuando las faltas principales eran la piratería, el maltrato de embajadores y la violación de conductas de seguridad. No obstante, al utilizar el término general "ley de naciones", el Congreso permitió que la cobertura del estatuto cambiara acorde a los cambios en el derecho internacional consuetudinario.

Al fallar a favor de Firestone, mi corte destacó dos cosas. La primera fue que, aunque el derecho estadunidense es muy estricto con las responsabilidades corporativas cuando se trata de agravios o cuasidelitos en los que incurren sus empleados en el curso de su trabajo, según la teoría de que esta responsabilidad estricta por los agravios de sus empleados provee a las corporaciones de incentivos para regular los que son necesarios porque, por lo general, los empleados son inmunes a los juicios y, por tanto, no están sujetos a sanciones por agravio, la misma teoría se debilita cuando los empleados incluyen a habitantes de naciones tercermundistas. Las empresas estadunidenses con subsidiarias en estos países suelen verse incapacitadas para prevenir los abusos cometidos contra los trabajadores por parte de los supervisores locales.[18]

[15] 28 U.S.C., 1350.

[16] Flomo *vs.* Firestone Natural Rubber Co., LLC, 643 F.3d, 1013 (7th Cir., 2011).

[17] The Paquete Habana, 175 U.S., 677, 700 (1900).

[18] Véase también Alan O. Sykes, "Corporate Liability for Extraterritorial Torts under the Alien Tort Statute and Beyond: An Economic Analysis", 100 *Geo. L. J.*, 2161 (2012).

La segunda fue que no se nos había proporcionado información acerca de la situación de los niños liberianos que no viven (y, por tanto, trabajan) en la plantación de Firestone. Sin embargo, lo que se sabía era que a los padres de los niños de la plantación se les pagaba muy bien para el promedio en Liberia. Asimismo, esto ponía a los niños que ayudaban a sus padres en mejores condiciones que a la mayoría de los niños en Liberia a pesar de tener que recolectar caucho, y podrían estar peor si sus padres, al ser incapaces de cumplir las cuotas diarias, perdieran sus trabajos o tuvieran que pagar a ayudantes adultos, reduciendo de esta manera los ingresos familiares. Existe un trueque de ganancias y sacrificios entre el ingreso familiar y el trabajo infantil: los niños se ven beneficiados por el primero y dañados por el segundo; el efecto neto sobre su bienestar no es muy claro.

XXXI.6. Las cláusulas de elección de foro o de jurisdicción

La decisión acerca de qué tribunal se encargará de litigar una demanda es particularmente importante cuando la elección no es entre estados o distritos federales estadunidenses, sino entre tribunales de países distintos. Por ello, las partes de un contrato internacional por lo general incluyen una cláusula de elección de foro o de jurisdicción (selección de un tribunal) en el contrato. Otro caso reciente servirá para ejemplificar lo complejo de esto.[19]

Los demandantes habían comprado tiempos compartidos en un centro turístico en México. Cada uno de los contratos entre el comprador y el vendedor contenía una cláusula que indicaba (en un inglés deficiente) que, "en caso de existir alguna controversia en la interpretación y cumplimiento de los derechos y obligaciones de este acuerdo, ambas partes acceden por la presente a someterse a las leyes aplicables y a los tribunales competentes de la ciudad de México, Distrito Federal, renunciando al derecho de elegir otro foro que les corresponda por efecto de su domicilio actual o a futuro". La demanda, llevada al tribunal federal de distrito en Chicago, alegaba que los acusados, Raintree y Starwood, habían defraudado a los demandantes con la venta de tiempos compartidos. Había sido una subsidiaria de Raintree, y no Raintree, la que había realizado la venta, pero la subsidiaria ya no existía, razón por la que la demanda se hacía en contra de Raintree, la matriz. Empero, Raintree, a pesar de no ser una de las partes del contrato (ni lo era tampoco Starwood, pero de eso hablaremos más adelante), apeló a la cláusula de elección de foro o de jurisdicción porque prefería litigar en un tribunal en México. Los demandantes habían acordado que cualquier disputa con el vendedor de los tiempos compartidos se resolviera en México, por lo que no les afectó que la matriz tomara el lugar de la subsidiaria.

[19] Adams *vs.* Raintree Vacation Exchange, LLC, 702 F.3d, 436 (7th Cir., 2012).

El enfoque literal a la hora de interpretar las cláusulas de elección de foro o de jurisdicción (un enfoque que nunca permitiría la sustitución de afiliados de los firmantes) podía socavar la contribución a la certeza respecto a las transacciones comerciales que se les ha reconocido a estas cláusulas, en particular en lo referente a las transacciones internacionales. Este enfoque crea certeza respecto a las partes que pueden apelar a esta cláusula (sólo los firmantes), pero incertidumbre respecto a la jurisdicción, pues una de las partes puede ser capaz de evitar el foro designado al manipular la relación de los afiliados. Por ejemplo, el firmante de un contrato con esta cláusula podría sustituir a la empresa sujeta al contrato con una corporación afiliada, quizás una creada con el único propósito de dar residencia nueva a la empresa, anulando la cláusula en el proceso. Así, un firmante que quisiera aplicar la cláusula se vería inhabilitado de cambiar su empresa a un afiliado corporativo a pesar de que el cambio fuera sensato desde el punto de vista comercial. Es mejor permitir que las partes decidan en el contrato si limitar la cláusula de elección de foro o de jurisdicción sólo a las entidades mencionadas a que la ley imponga estos límites como una disposición por defecto en caso de no existir una especificación respecto a las otras entidades sujetas al contrato. Este segundo enfoque complicaría las negociaciones de estas cláusulas porque las partes tendrían que esforzarse más en cerrar todas las fisuras y lagunas que se abrirían si tan sólo las entidades mencionadas en el contrato pudieran *alguna vez* apelar o ser sujetas a esta cláusula.

Starwood, a diferencia de Raintree, no era parte de la cadena de titulares de los demandantes. Los demandantes aseguraban que ambas empresas habían conspirado para que la subsidiaria de Raintree utilizara el dinero de la venta de tiempos compartidos para pagar una deuda que Raintree le debía a Starwood, en lugar de construir las villas. La cuestión era si un presunto conspirador podía apelar a la cláusula de elección de foro o de jurisdicción en un contrato, firmado por su presunto coconspirador, que creaba o fomentaba dicha conspiración. La respuesta era sí. Se acusaba a Raintree y a Starwood de ser los principales secretos de la subsidiaria de Raintree, a su vez agente que trataba con los compradores de los tiempos compartidos y, por tanto, efectuaba el fraude. Un contrato que el agente de los principales secretos hace con una tercera parte puede hacerse cumplir, por decisión de la tercera parte desde luego (la tercera parte puede rescindirlo si prefiere), contra los principales secretos. De aquí se seguía que, en la situación contraria, Starwood podría someter a los demandantes a la cláusula y así defender la demanda en México. De no ser por este principio de reciprocidad, los demandantes tendrían la opción de elegir el foro o la jurisdicción, pero Starwood no. Ésta no podría haber sido la intención subyacente a la cláusula que establecía a México como la jurisdicción exclusiva independientemente del domicilio de las partes involucradas.

Desde un punto de vista práctico, era obvio que el caso tenía que litigarse como uno solo en un solo tribunal de un solo país, en lugar de dos casos en dos

tribunales en dos países. Supongamos que los demandantes hubiesen querido demandar a Raintree y a Starwood en México. Estarían autorizados a demandar a Raintree ahí sin importar la reciprocidad; ¿podría Starwood insistir en que se *le* demandara en los Estados Unidos? Esta división en el juicio sería muy ineficiente. Y notemos que, debido a que Raintree tenía la autoridad de llevar el caso a México de acuerdo con la cláusula de elección de foro o de jurisdicción independientemente de los derechos de Starwood, el tribunal de los Estados Unidos también podría revocar la demanda contra Starwood aun si no tuviera la autoridad para exigir el cumplimiento de la cláusula de elección de foro o de jurisdicción. Esta desestimación procedería por el simple hecho de que, al realizarse la demanda en México en contra del coacusado, el tribunal estadunidense sería una jurisdicción inconveniente para una demanda paralela en contra de Starwood.

Lecturas recomendadas

Bussani, Mauro, Veron Valentine Palmer y Francesco Parisi, "Liability for Pure Financial Loss in Europe: An Economic Restatement", 51 *Am. J. Comp. L.*, 113 (2003).

Carbonara, Emanuela y Francesco Parisi, "Choice of Law and Legal Evolution: Rethinking the Market for Legal Rules", 139 *Pub. Choice*, 461 (2009).

Dunoff, Jeffrey L. y Joel P. Trachtman, "Economic Analysis of International Law", 24 *Yale J. Int'l L.*, 1 (1999).

Economic Dimensions in International Law: Comparative and Empirical Perspectives (Jagdeep S. Bhandari y Alan O. Sykes [comps.], 1997).

Economics of International Trade Law (2 vols.; Alan O. Sykes [comp.], 2012).

Goldsmith, Jack L. y Eric A. Posner, *The Limits of International Law* (2005).

Kalajdzic, Jasminka, Peter Cashman y Alana Longmoore, "Justice for Profit: A Comparative Analysis of Australian, Canadian and U.S. Third Party Litigation Funding", 61 *Am. J. Comp. L.*, 93 (2013).

Kovač, Mitja, *Comparative Contract Law and Economics* (2011).

Ordover, J. A., A. O. Sykes y R. D. Willig, "Unfair International Trade Practices", 15 *N.Y.U. J. Int'l Law & Pol.*, 323 (1983).

Posner, Eric A. y Alan O. Sykes, *Economic Foundations of International Law* (2013).

Sykes, Alan O., "Public *versus* Private Enforcement of International Economic Law: Standing and Remedy", 34 *J. Legal Stud.,* 631 (2005).

——— "Countervailing Duty Law: An Economic Perspective", 89 *Colum. L. Rev.*, 199 (1989).

"Symposium on Rational Choice and International Law", 31 *J. Legal Stud.,* S1 (2002).

Problemas

1. Cuando se utiliza gas venenoso para la guerra, por lo general se encapsula en un proyectil de artillería que se rompe al tocar el suelo liberando el gas. En ese caso, ¿por qué está prohibido el uso de gases venenosos como armas de guerra, pero no el uso de explosivos de alta potencia en proyectiles de artillería?

2. ¿Debiera ser posible recurrir al Estatuto de Reclamación por Agravios contra Extranjeros en un caso en el que los demandantes son ciudadanos extranjeros, el acusado es un ciudadano extranjero y el atentado contra "la ley de las naciones o un tratado de los Estados Unidos" ocurrió en su totalidad en una nación extranjera?

ÍNDICE ANALÍTICO

abandono de propiedad: 59, III.2; 113, III.13
aborto: 209, V.4-220, V.7; derecho a: 1040, XXV.3
adicción; razón: 339, VII.11; impuesto: 775, XVIII.3
adjudicación: 847, XX.2; 914, XXII.7; 951, XXII.16; anticipatoria: 900, XXII.3; óptima: 953, XXII.17. *Véase también* jueces
administradores fiduciarios: 152, IV.8; 668, XV.14. *Véanse también* agencia, costos de; agencia: derecho de; fideicomisos
adopción: 209, V.4; 215, V.5. *Véase también* venta de bebés
adulterio: 199, V.2
agencia administrativa o pública: 996, XXIV.1-1009, 24.4; 1016, XXIV.6-1025, 24.9; derecho de: 152, IV.8; revisión judicial de decisiones: 1027, XXIV.10
agencia, costo: 615, XV.1; 633, XV.5; 646, XV.7; 650, XV.8; 1025, XXIV.9; judicial: 860, XX.6
agencias administrativas y derecho administrativo: 568, XIV.1; 573, XIV.2; 886, XXI.3; 956, XXII.18; 996, XXIV.1; 1009, XXIV.4
agencias, superposición de: 1014, XXIV.5
Alaska Packers' Assn. *vs*. Domenico: 128, IV.2
altruismo: 153, IV.9; 184, IV.15; 264, VI.9; en empresas de propiedad familiar: 370, VIII.6; dentro de la familia: 120, IV.1; 194, V.1; impulso para el rescate: 264, VI.9; motivación de la donación: 791, XVIII.8; motivación de las donaciones: 816, XIX.1
American Commercial Lines *vs*. Louisville & N.R.R.: 548, XIII.5
Amor sin barreras (Robbins y Bernstein): 481, XI.5
análisis costo-beneficio: 593, XIV.9; aplicado a los casos de libertad de expresión: 1115, XXIX.4; del calentamiento global: 596, XIV.10; en casos constitucionales: 220, V.7; de la contaminación: 584, XIV.6; como herramienta reguladora: 593, XIV.9; del interrogatorio coercitivo: 1134, XXX.3; de riesgos catastróficos: 596, XIV.10; en sentido inverso: 596, XIV.10. *Véase también* descuento al valor presente
análisis económico del derecho: 45, II.1; 50, II.4; 886, XXI.3; positivo *vs*. normativo: 47, II.2
análisis *ex ante vs*. análisis *ex post*: 15, I.1; 273, VI.12; regulación: 317, VII.3; 331, VII.7; 568, XIV.1; 1009, XXIV.4; 1113, XXIX.3
anticoncepción: 217, V.6; 220, V.7; 1122, XXIX.7
antiseguro: 159, IV.10
apelaciones: 956, XXII.18; costo de: 910, XXII.6; de las decisiones administrativas: 1004, XXIV.3; 1027, XXIV.10; doctrina del error inocuo: 985, XXIII.5; interlocutoria: 956, XXII.18; tasa: 928, XXII.12
arbitraje: 839, XX.1; 953, XXII.17; en la Islandia medieval: 378, VIII.8; obligatorio: 506, XII.3
Aristóteles: 49, II.3; 375, VIII.7; 973, XXIII.2
arresto, *véase* cateos y confiscaciones
arte: 100, III.11
asesoría, derecho a la: 1040, XXV.3
asociaciones de casatenientes: 86, III.8
ataduras: 444, X.10
autoridad aparente: 152, IV.8; 190, cap. IV, prob. 11
Ayuda para Familias con Hijos Dependientes (AFDC): 751, XVII.6

banca: 720, XVI.14-730; XVI.16; crisis, 730, XVI.14; regulación de, 716, XVI.13; 727, XVI.15
Beccaria, Cesare: 45, II.1
Becker, Gary S.: 15, I.1; 45, II.1; 209, V.4; 217, V.6; 339, VII.11; 370, VIII.6; 509, XII.5; 746, XVII.4; 751, XVII.6; 769, XVIII.1; 775, XVIII.3; 850, XX.3; 867, XX.7; 996, XXIV.1; 1016, XXIV.6; 1086, XXVIII.1
Bentham, Jeremy: 15, I.1; 45, II.1
Bentley *vs*. State: 120, IV.1
Bird *vs*. Holbrook: 285, VI.15
Blackstone, William: 56, III.1
blockbusting: 1127, cap. XXIX, prob. 1

blogs: sus derechos de cara a la Primera Enmienda: 1127, cap. XXIX, prob. 7
Blyth *vs*. Birmingham Water Works: 229, VI.1
boicots: 548, X.12
Boomer *vs*. Atlantic Cement Co.: 91, III.9
Brandeis, Louis D.: 1047, XXVI.1
Brown *vs*. el Departamento de Educación: 1088, XXVIII.2; 1091, XXVIII.3
Brunswick Corp. *vs*. Pueblo Bowl-O-Mat, Inc.: 461, X.14
buena voluntad: 132, IV.3; 668, XV.14

calentamiento global: 217, V.6; 596, XIV.10; análisis costo-beneficio de: 596, XIV.10; y Ley del Aire Limpio: 596, XIV.6; Protocolo de Kioto: 596, XIV.10
Cámara de Comercio *vs.* Dow Jones & Co.: 495, cap. XI, prob. 6
canibalismo: 332, VII.8
capacidad excedente: 253, VI.7; 410, X.1; 438, X.9; 565, cap. XIII, prob. 2; *vs.* capacidad de "carga máxima": 253, VI.7
capital humano: 208, V.4; 273, VI.12; 301, VII.1; 306, VII.2; 515, XII.8; 521, XII.10; 531, XII.11; 615, XV.1; 658, XV.11; 730, XVI.16; 744, XVII.3; 746, XVII.4; 783, XVIII.5; específico de la empresa: 400, IX.4; 410, X.1; 501, XII.2; 507, XII.4; 509, XII.5; 515, XII.8; 521, XII.10; 531, XII.11; 628, XV.4; 633, XV.5; general *vs*. específico de la empresa: 509, XII.5; del mercado *vs*. hogareño: 521, XII.10; donaciones de: 816, XIX.1
capital, costo del: 267, VI.11; 540, XIII.2; 543, XIII.3; 689, XVI.2; 701, XVI.7; 783, XVIII.5; 797, XVIII.9; estructura de la empresa regulada: 537, XIII.3. *Véanse también* corporaciones; capital humano; ganancias de capital; tributación
carga de la prueba: 979, XXIII.4; 994, cap. XXIII, prob. 3; en caso de negligencia: cap. 994, XXIII, prob. 2; civil *vs*. penal: 979, XXIII.4; del empleador: 521, XII.10; más allá de una duda razonable: 979, XXIII.4
carga de trabajo, judicial: 860, XX.6; 947, XXII.15; 956, XXII.18; 1042, XXV.4
caridad: 359, VIII.2; corporativa: 668, XV.14; "prodigada para hacer arrancar a la economía": 668, XV.14
cárteles, *véase* derecho antimonopólico; monopolio
castigo colectivo: 260, VI.8; 317, VII.3; 366, VIII.5
cateos y confiscaciones: 1128, XXX.1-1134, XXX.3; bloqueos al camino como una forma de: 1128, XXX.1; cateos aleatorios: 1128, XXX.1; inmunidad oficial: 1131, XXX.2; parada de Terry: 1128, XXX.1; regla excluyente: 1131, XXX.2; remedios del cuasidelito: 1131, XXX.2
cercas; 68, III.5; 108, III.12
chantaje: 330, VII.6; 999, XXIV.2; 1105, XXIX.2
cláusulas de elección de foro o de jurisdicción: 1156, XXXI.6
coacción: 153, IV.9; 172, IV.12. *Véase también* bajo contratos
Coase, Ronald H.: 15, I.1; 39, I.5; 45, II.1; 68, III.5; 407, IX.6; 548, XIII.5; 615, XV.1; 1119, XXIX.6
colusión, tácita *vs*. explícita, *véase* derecho antimonopólico
comercio, interestatal: 1057, XXVII.1; 1065, XXVII.4; 1073, XXVII.6; 1078, XXVII.10; 1080, XXVII.11; 1083, cap. XXVII, prob. 9; 1093, XXVIII.4
Comisión Federal de Comercio: 568, XIV.1-580, XIV.3; 1014, XXIV.5-1020, XXIV.7; 1119, XXIX.6
Comisión Federal de Comunicaciones: 64, III.3; 603, XIV.11; 1016, XXIV.6; 1020, XXIV.7; 1117, XXIX.5
Comisión Federal de Energéticos: 565, cap. XIII, prob. 9
Comisión Federal de Radio: 64, III.3; 1016, XXIV.6
Comisión Federal Reguladora de la Energía: 565, cap. XIII, prob. 9; 1068, XXVII.4
compensación: contratos de tenencia: 1054, XXVI.3; diferida o demorada: 521, XII.10; 720, XVI.14; judicial: 847, XX.2. *Véase también* derecho del empleo; capital humano; mercados de trabajo; salarios
competencia: entre abogados: 962, XXII.20; como cuasidelito: 253, VI.7; 603, XIV.11; 839, XX.1; efecto de, sobre la discriminación racial: 1086, XXVIII.1; entre los estados para atraer a las corporaciones: 652, XV.10; fuera de los precios: 396, IX.3; monopólica: 159, IV.10; por el monopolio o por el mercado: 407, IX.6; 596, XIV.10; per-

fecta: 391, IX.1; 444, X.10. *Véase también* monopolio
complementariedad: 1009, XXIV.4
confesiones: 886, XXI.3; 1134, XXX.3. *Véase también* autoincriminación
conflicto de leyes, *véase* elección de la ley
conflictos de interés, *véase* agencia, costos de; profesión de abogado
Congreso, como junta de directivos del ramo ejecutivo: 1036, XXV.2
conscripción: 773, XVIII.2; de jurados: 951, XXII.16
Conseil d'État: 370, VIII.6
consideración: 128, IV.2; existencia *vs*. adecuación: 128, IV.2; moral: 128, IV.2
conspiración: 317, VII.3; 410, X.1. *Véase también* derecho antimonopólico
Constitución: 1033, XXV.1; alcance del poder del Congreso para regular el comercio: 1057, XXVII.1; 1080, XXVII.11; analogía contractual: 1033, XXV.1; el caso de los bonos escolares: 1123, XXIX.8; casos de los Diez Mandamientos: 1023, XXIX.8; cláusula de comercio: 1057, XXVII.1; 1080, XXVII.11; constitucionalidad de las agencias administrativas: 1036, XXV.2; derecho de portar armas (Segunda Enmienda): 380, VII.9; derechos constitucionales de sospechosos de terrorismo: 1042, XXV.4; doctrina del participante en el mercado: 1083, cap. XXVII, prob. 6; límites de daños punitivos: 285, VI.15; un método de redistribución de la riqueza: 1040, XXV.3; requisito de la acción estatal en la Decimocuarta Enmienda: 1091, XXVIII.3; revisión de la racionalidad: 1044, XXV.5; separación de poderes: 1036, XXV.2; temas económicos relacionados con: 1033, XXV.1. *Véase también federalismo; y bajo disposiciones y doctrinas constitucionales particulares*
construir una casa en tierras públicas: 66, III.4; 488, XI.7
contadores: 152, IV.8
contratos: de adhesión: 153, IV.9; buena fe en el cumplimiento: 120, IV.1; 152, IV.8; 185, IV.16; cláusulas de castigo: 172, IV.12; 190, cap. IV, prob. 11; 1141, XXXI.1; cláusulas de "tómalo o paga": 172, IV.12; cláusulas del mejor esfuerzo: 120, IV.1; coacción: 153, IV.9; 172, IV.12; confianza: 159, IV.10; 190, cap. IV, probs. 3, 9, 12 y 13; cumplimiento sustancial: 179, IV.14; doctrinas de la imposibilidad y relacionadas: 137, IV.5; 899, XXII.2; ejecutorio: 120, IV.1; 159, IV.10; 160, cap. IV, prob. 8; empleo: 501, XII.2; 509, XII.5; *vs*. empresa como método de organización de la producción: 615, XV.1; error mutuo: 134, IV.4; entre estados: 1033, XXV.1; estándares impresos: 153, IV.9; 623, XV.3 (*véase también* de adhesión); como una forma de seguro: 137, IV.5; de futuros *vs*. a futuro: 159, IV.10; garantía implicada de la adecuación para el uso intentado: 137, IV.5; gobierno: 190, cap. IV, prob. 11; 509, XII.5; 568, XXIV.1; ilegalidad de, como defensa: 190, cap. IV, prob. 8; inconsciencia: 153, IV.9; incumplimiento eficiente: 120, IV.1; 153, IV.9; 172, IV.12; 185, IV.16; 356, VIII.2; 1141, XXXI.1; interpretación de: 120, IV.1; 142, IV.6; de largo plazo: 407, IX.6; 603, XIV.11; 1033, XXV.1; modificación: 128, IV.2; 172, IV.12; multilateral *vs*. bilateral: 1144, XXXI.2; ofrecimiento y aceptación: 132, IV.3; 153, IV.9; promesa de incumplimiento colateral: 190, cap. IV, prob. 9; de requerimientos: 407, IX.6; responsabilidad estricta por el incumplimiento de: 273, IV.12; 292, IV.16; 1144, XXXI.2; de seguro: 142, IV.6; para servicios personales: 137, IV.5; 199, V.2; teoría de la opción de, o de las obligaciones contractuales de Holmes: 172, IV.12; 185, IV.16; 1141, XXXI.1; teoría objetiva *vs*. teoría "subjetiva": 132, IV.3; tratados como: 1144, XXXI.2; unilaterales: 132, IV.3
cooperativas de trabajadores: 618, XV.2; 702, XVI.8. *Véase también* ESOP (Plan de Propiedad de Acciones de los Empleados)
coordinación por planificadores centrales *vs*. por mercados: 615, XV.1
copyright: abandono de: 468, XI.1; 495, cap. XI, prob. 5; licencia obligatoria: 185, IV.16; 603, XIV.11; prob. de la duración: 481, XI.5; software de computadora: 410, X.1; televisión por cable y: 603, XIV.11; 612, cap. XIV, probs. 2 y 3; violación o infracción contributiva: 481, XI.5; 612, cap. XIV, prob. 3
corporaciones (empresas): acreedores asegurados *vs*. no asegurados: 628, XV.4; 633, XV.5; cerradas *vs.* públicas: 646, XV.7; 676, XV.16; compra de seguro por: 692, XVI.3;

conglomeradas: 689, XVI.2; Delaware: 656, XV.10; democracia corporativa: 652, XV.9; 867, XX.7; descorrer el velo corporativo: 646, XV.7; discriminación racial por: 680, cap. XV, prob. 9; 1086, XXVIII.1; donaciones corporativas: 668, XV.14; efecto de la competencia estatal: 656, XV.10; expulsiones: 656, XV.10; juicios derivados: 668, XV.14; lucha de poderes: 652, XV.9; 704, XVI.9; maximización de la ganancia *vs.* maximización de las ventas o el crecimiento: 668, XV.14; negocios de funcionarios internos: 658, XV.11; 680, cap. XV, prob. 7; sin fines de lucro: 618, XV.2; obligaciones de lealtad y cuidado de los administradores: 650, XV.8; oferta de compra: 652, XV.9; opciones de acciones: 658, XV.11; 664, XV.13; 680, cap. XV, prob. 3; prima por el control: 650, XV.8; 652, XV.9; responsabilidad limitada de los accionistas: 26, I.2; 294, VI.17; 623, XV.3; 628, XV.4; 633, XV.5; 646, XV.7; responsabilidad penal de: 674, XV.15; responsabilidad social de: 668, XV.14; 674, XV.15; separación de la propiedad y el control: 650, XV.8; 687, XX.7; transferencia del control corporativo: 652, XV.9; tributación: 783, XVIII.5; 1071, XXVII.5. *Véase también* agencia, costos de; distinción entre la empresa y la corporación; valores; valores, regulación de los

corrupción: 370, VIII.6; 608, XIV.12

costos de la información: 15, I.1; 94, III.10; 209, V.4; 245, VI.5; 285, VI.15; privados: 245, VI.5; para el consumidor: 153, IV.9; y discriminación inversa: 1096, XXVIII.5; como un factor en la discriminación: 521, XII.10; influencia de, sobre el derecho primitivo: 366, VIII.5; e intención criminal: 326, VII.5; de los jueces: 886, XXI.3; 893, XXI.4; 1044, XXV.4; para los jueces *vs.* para los legisladores y los funcionarios de la rama ejecutiva: 855, XX.4; 1042, XXV.4; en el proceso legislativo: 886, XXI.3; de reglas contra estándares: 886, XXI.3. *Véase también* vaguedad

costos: comunes: 410, X.1; 548, XIII.5; conjuntos: 548, XIII.5; 1123, XXIX.8; de corto plazo *vs*. de largo plazo: 436, X.8; 438, X.9; 548, XIII.5; encallados: 558, XIII.9; fijo: 159, IV.10; 172, IV.12; 370, VIII.6; 410, X.1; 438, X.9; 537, XIII.1; 548, XIII.5; 633, XV.5; hundidos: 15, I.1; 33, I.4; 468, XI.1; marginales: 15, I.1, n 9; 113, III.13; 159, IV.10; 245, VI.5; 251, VI.6; 391, IX.1-391, IX.3; 409, cap. IX, prob. 1; 410, X.1; 423, X.3; 426, X.4; 430, X.6; 438, X.9; 458, X.12; 490, XI.8; 537, XIII.1; 548, XIII.5; 554, XIII.6; 584, XIV.6; medios totales: 159, IV.10; 410, X.1; 438, X.9; 455, X.11; 812, cap. XVIII, prob. 11; medios variables: 159, IV.10; 438, X.9; 548, XIII.5; de oportunidad: 15, I.1; 267, VI.11; 391, IX.1; 400, IX.4; 773, XVIII.2; originales *vs*. de remplazo: 444, X.10; 540, XIII.2; 543, XIII.3; variables: 159, IV.10; 410, X.1; 438, X.9; 444, X.10; 633, XV.5; venta por debajo del: 410, X.1. *Véase también* costos de transacción

costumbre: 364, VIII.4; 378, VIII.8

Cuarta Enmienda, *véase* cateos y confiscaciones

daños: antimonopólicos: 461, X.14; en los casos de cateos y confiscaciones: 1131, XXX.2; 1134, XXX.3; en los casos de muerte: 351, VIII.1; 273, VI.12; en los casos de discriminación: 1093, XXVIII.4; en los casos de valores: 356, VIII.2; 603, XVI.11; cálculos de las ganancias futuras perdidas: 267, VI.11; consecuenciales: 170, IV.11; 170, IV.11; contractuales: 159, IV.10; cuasidelitos: 237, VI.4; dolor y sufrimiento: 273, VI.12; para el empleador cuando un empleado muere o es herido: 273, VI.12; fallo o arreglo de pago triple: 410, X.1; 461, X.14; 968, cap. XXII, prob. 7; hedonísticos: 273, VI.12; inmunidad oficial: 296, VI.18; por pérdida de la producción hogareña: 204, V.3; 267, VI.11; previsibilidad de: 253, IV.7; *vs*. prohibiciones: 91, III.9; punitivos: 172, IV.12; 253, VI.7; 285, VI.15; 301, VII.1; 356, VIII.2; 461, X.14; 674, XV.15; 689, XVI.2; regla de los beneficios colaterales (fuente colateral): 278, VI.13; 279, VI.14; como remedio privado implicado en el estatuto regulador: 855, XX.4; enfoque fiscal de: 584, XIV.6; 791, XVIII.8

debido proceso legal: 897, XXII.1; 1040, XXV.3; 1044, XXV.5; 1047, XXVI.1-1054, XXVI.3; 1062, XXVII.3; 1065, XXVII.4; 1074, XXVII.7; 1091, XXVIII.3; 1123, XXIX.8

Decimocuarta Enmienda, *véanse* debido proceso legal; discriminación racial; protección igual de las leyes

delitos y derecho penal: y castigo: 301, VII.1-306, VII.2; 344, VII.12; castigo colectivo: 260, VI.8; 317, VII.3; 366, VIII.5; causación: 253, VI.7; comparados con los cuasidelitos intencionales: 285, VI.15; 301, VII.1; consentimiento como defensa: 301, VII.1; 348, cap. VII, prob. 9; conspiración: 317, VII.3; corporativos: 674, XV.15; crimen organizado: 336, VII.10; de descuido: 331, VII.7; defensa de locura: 326, VII.5; defensa de necesidad: 306, VII.2; 332, VII.8; derecho común: 317, VII.3; en las sociedades primitivas: 366, VIII.5; disuasión *vs.* incapacitación: 306, VII.2; 326, VII.5; 344, VII.12; disuasión marginal: 306, VII.2; 317, VII.3; 331, VII.7; doctrina Pinkerton: 317, VII.3; drogas: 339, VII.11; efecto de los antecedentes penales sobre la instrucción: 336, VII.10; 979, XXIII.4; 988, XXIII.7; efecto estigmatizante de la condena: 306, VII.2; 317, VII.3; elasticidad de la oferta de delincuentes: 339, VII.11; entrampamiento: 317, VII.3; 348, cap. VII, prob. 10; escuelas para delincuentes: 996, XXIV.1; experimentos médicos con prisioneros: 50, II.4; extorsión: 336, VII.10; fraude bancario *vs*. robo de banco: 1062, XXVII.3; heredero asesino: 823, XIX.2; homicidio con delito: 317, VII.3; 326, VII.5; impulsivos *vs.* deliberados: 326, VII.5; incitación para cometer: 1105, XXIX.2; incoados: 204, VII.3; infractores habituales: 988, XXIII.7; instigación y complicidad: 317, VII.3; intención específica *vs*. general: 326, VII.5; jurisdicción federal *vs*. estatal sobre: 1062, XXVII.3; leyes de "tres *strikes*": 306, VII.2; multas *vs*. prisión: 306, VII.2; negligentes: 823, XIX.2; de obscenidad: 1112, XXIX.3; 1127, cap. XXIX, prob. 4; de odio: 330, VII.6; penas de vergüenza: 306, VII.2; perdones: 325, VII.4; perfil étnico o racial: 1096, XXVIII.5; pobreza como causa de: 348, cap. VII, prob. 8; 751, XVII.6; premeditados *(mens rea)*: 326, VII.5; prevención *vs.* castigo: 317, VII.3; 326, VII.5; 344, VII.12; prisión: 344, VII.2; prosecución y fiscales: 918, XXII.9; 979, XXIII.4; 1004, XXIV.3; 1025, XXIV.9; prostitución: 336, VII.10; provocación como defensa parcial: 334, VII.9; reincidencia: 317, VII.3; responsabilidad estricta: 326, VII.5; 331, VII.7; riesgo de error en el castigo penal: 306, VII.2; sanciones penales óptimas: 306, VII.2; sentencia: 306, VII.2; 334, VII.9; 919, XXII.10; sexuales anteriores: 988, XXIII.7; invitación a cometer: 1105, XXIX.2; tasas delictivas: 346, VII.13; 979, XXIII.4; tentativas: 317, VII.3; vaguedad de las leyes penales: 886, XXI.3; víctimas: 301, VII.1; 306, VII.2; 330, VII.6; 331, VII.7; 334, VII.9. *Véase también* aplicación forzosa de la ley; autoincriminación; evidencia (pruebas); leyes de control de armas; procedimiento: penal

demanda: derivada: 458, X.12; elasticidad cruzada: 432, X.7; 769, XVIII.1. *Véase también* elasticidad de la demanda *y* Ley de la Demanda

demandas colectivas: 928, XXII.12

democracia: 867, XX.7; 877, XX.9; 1033, XXV.1; corporativa: 652, XV.9; 867, XX.7; deliberativa: 1127, cap. XXIX, prob. 5; directa: 877, XX.9; grupos de interés: 850, XX.3; teoría del votante mediano: 740, XVII.2

demografía: 217, V.6

demora: 137, IV.5; 172, IV.12

depreciación: 68, III.5; 306, VII.2; 410, X.1; 540, XIII.2; 783, XVIII.5; 786, XVIII.6; 882, XXI.1

depresión, económica: 172, IV.12; 506, XII.3; 509, XII.5; 684, XVI.1; 710, XVI.11; 727, XVI.15; 789, XVIII.7; 1047, XXVI.1; 1150, XXXI.3

derecho antiguo, *véase* derecho primitivo y antiguo

derecho antimonopólico: 410, X.1-461, X.14; 466, cap. X, probs. 4, 6; 514, XII.7, 537, XIII.1; aplicación de: 1014, XXIV.5; comparado: 1141, XXXI.1; daños: 411, X.1; 461, X.14; y la economía del federalismo: 1083, cap. XXVII, probs. 7, 8; su enfoque de la democracia: 877, XX.9; interacción entre, y propiedad intelectual: 490, XI.8

derecho comparado: 1141, XXXI.1. *Véase también* derecho: y modernización; derecho: primitivo y antiguo

derecho común: como costumbre: 364, VIII.4; 378, VIII.8; eficiencia de: 351, VIII.1; 380, VIII.9; 839, XX.1; 847, XX.2; 860, XX.6; 882, XXI.1; 914, XXII.7; *vs.* la legislación: 839, XX.1; moralidad de: 380, VIII.9; papel del precedente en: 893, XXI.4; 914, XXII.7; papel en el crecimiento económico: 359,

VIII.3; 839, XX.1; *vs.* la regulación: 568, XIV.1; relación con la justicia correctiva: 49, II.3; relación con la justicia distributiva: 380, VIII.9. *Véase también bajo ramas específicas del derecho común*
derecho del empleo: acreditación ocupacional: 962, XXII.20; antidiscriminación: 521, XII.9; 521, XII.10; 1093, XXVIII.4; contrato de tenencia como propiedad constitucional: 1054, XXVI.3; derecho de amo-sirviente: 615, XV.1; empleo a voluntad: 120, IV.1; 509, XII.5; 521, XII.10; 1054, XXVI.3; federalización de: 1062, XXVII.3; leyes de: 501, XII.2; leyes del salario mínimo y de la jornada laboral máxima: 515, XII.8; 847, XX.2; 1047, XXVI.1; 1056, cap. XXVI, prob. 2; leyes del trabajo infantil: 194, V.1; 199, V.2; 209, V.4; 515, XII.8; 674, XV.15; negociación colectiva: 497, II.2-509, XII.5; 514, XII.7; 515, XII.8; 356, cap. XII, prob. 1; regulación de la seguridad y la salud del trabajador: 515, XII.8; salario de eficiencia: 521, XII.10; términos obligatorios: 521, XII.10; 531, XII.11; tratamiento de los sindicatos en el derecho común: 497, XII.1. *Véase también* compensación de los trabajadores; empleo; ERISA; mercados de trabajo
derecho familiar y economía: 194, V.1-224, V.8; acuerdo familiar: 209, V.4; beneficios obligatorios de la maternidad: 521, XII.9; y homosexualidad: 220, V.7; 224, V.8. *Véase también* aborto; divorcio; hijos; matrimonio; política poblacional; producción hogareña
derecho internacional: Estatuto de Reclamación por Agravios contra Extranjeros: 1155, XXXI.5; *antidumping* y cuotas compensatorios: 1150, XXXI.3; inversión extranjera directa: 1153, XXXI.4; cláusulas de elección de foro o jurisdicción: 1156, XXXI.6; regulación de la pesca: 1144, XXXI.2; aranceles en venganza: 1150, XXXI.3; uso de la historia legislativa posterior a su proclamación: 855, XX.4
derecho marítimo: 153, IV.9; 237, VI.4; 351, VIII.1; 1144, XXXI.2
derecho primitivo y antiguo: 59, III.2; 74, III.6; 120. IV.1; 366, VIII.5; 996, XXIV.1
desacato: 159, IV.10; 178, IV.13; 185, IV.16; 301, VII.1; 968, cap. XXII, prob. 13; 1134, XXX.3
descuento: en el análisis costo-beneficio: 593, XIV.9; en los casos de cuasidelitos: 267, VI.11; 297, cap. VI, prob. 3; en los casos de libertad de expresión: 1105, XXIX.2; por delincuentes: 306, VII.2; hiperbólico: 33, I.4; y el problema del interés previo al juicio: 910, XXII.6; tasas de, privadas *vs.* sociales: 593, XIV.9; 835, cap. XIX, prob. 7; al valor presente: 33, I.4, n 15
deseconomías de escala: 159, IV.10; 336, VII.10; 537, XIII.1; 956, XXII.18; 1009, XXIV.4; 1057, XXVII.1. *Véase también* economías de escala
desregulación: 396, IX.3; 497, XII.1; 720, XVI.14; 727, XVI.15; efectos de riqueza de: 558, XIII.9; movimiento para: 47, II.2; 537, XIII.1; 558, XIII.9; 962, XXII.20; 1016, XXIV.6
difamación: 292, VI.16; limitaciones constitucionales a: 1105, XXIX.2
discapacidad: 267, VI.11; 273, VI.12; 297, cap. VI, prob. 3; 521, XII.10; 744, XVII.3; 751, XVII.6; 1027, XXIV.10
discriminación de precios: 405, IX.5; el agrupamiento como una forma de: 444, X.10; la atadura como una forma de: 444, X.10; atando el servicio de reparación o las piezas de repuesto al equipo original: 444, X.10; cuotas por derechos sobre la presentación: 947, XXII.15; depredadora: 438, X.9; descuento de fidelidad como una forma de: 444, X.10; distinta de las diferencias de precios relacionadas con los costos conjuntos u ordinaria: 548, XIII.5; inducida por la regulación: 548, XIII.5; y leyes *antidumping*: 1150, XXXI.3; perfecta (primer grado) *vs.* imperfecta (tercer grado): 405, IX.5; *vs.* precios de Ramsey: 548, XIII.5; selectiva *vs.* sistemática: 410, X.1
discriminación por la edad: 521, XII.10; 536, cap. XII, prob. 6
discriminación sexual: 194, V.1; 521, XII.10; acoso sexual: 224, VI.8; diferencias de salarios entre hombres y mujeres: 194, V.1; 521, XII.10; eficiente: XII.10; jornadas máximas y salarios mínimos para las mujeres: 1047, XXVI.1; Leyes de Paga Igual: 521, XII.10
discriminación: 50, II.4; 1086, XXVIII.1; 1100, cap. XXVIII, prob. 1; acción afirmativa: 521, XII.10; 1093, XXVIII.5; concepto del efecto disparejo: 521, XII.10; 1093, XXVIII.4; da-

ños por: 536, cap. XII, prob. 7; 1093, XXVIII.4; contra los discapacitados: 521, XII.10; efectos de riqueza de las leyes contra: 521, XII.10; en el empleo: 521, XII.10; 558, XIII.9; 1093, XXVIII.4; 1100, cap. XXVIII, prob. 7; 1096, estadística: XXVIII.5; de los estados en contra de no residentes: 1065, XXVII.4; 1083, cap. XXVII, prob. 6; leyes contra: 521, XII.10; 536, cap. XII, prob. 7; 1009, XXIV.4; 1093, XXVIII.4; 812, cap. XXVIII, prob. 7; contra la mayoría y no contra la minoría: 1040, XXV.3; perfiles étnicos en: 1096, XXVIII.5; inversa: 1040, XXV.3; 1096, XXVIII.5; 1100, cap. XXVIII, prob. 6. *Véase también* discriminación de precios; discriminación por la edad; discriminación por sexo; discriminación racial
distribución del ingreso y la riqueza, *véase* justicia distributiva; riqueza: distribución de
diversificación: 623, XV.3; 650, XV.8; 652, XV.9; 684, XVI.1; 689, XVI.2; 693, XVI.4; 951, XXII.16
diversificación: 623, XV.3; 650, XV.8; 652, XV.9; 684, XVI.1; 689, XVI.2; 693, XVI.4; 951, XXII.16
división del trabajo: 194, V.1; 217, V.6; 317, VII.3; 867, XX.7; 1036, XXV.2
divorcio: 15, I.1; 194, V.1; 199, V.2; 204, V.3; 301, VII.1; 755, XVII.7; 999, XXIV.2
doctrina Chenery: 1027, XXIV.10
doctrina de los compradores de buena fe: 100, III.11
dominio eminente: 74, III.6; 81, III.7; 116, cap. III, prob. 13; como remedio del ruido de los aviones: 81, III.7; compensación de las mejoras: 116, cap. III, prob. 4; confiscaciones reguladoras: 74, III.6; intereses divididos: 116, cap. III, prob. 4; litigación: 74, III.6; como método de tributación: 74, III.6; privado: 74, III.6; requerimiento de "uso público": 74, III.6
donaciones: 816, XIX.1; 829, XIX.4; 791, XXVIII.8; de capital humano: 816, XIX.1; tributación de: 769, XVIII.1; 791, XVIII.8; 816, XIX.1; 835, cap. XIX, probs. 1, 5 y 6
Dr. Miles Medical Co. *vs*. John D. Park & Sons Co.: 426, X.4
droit de suite: 495, cap. XI, prob. 3
dumping: 1150, XXXI.3
duopolio: 432, X.7; 490, XI.8; político: 867, XX.7; 877, XX.9
Eastern States Retail Lumber Dealers' Assn. *vs.* United States: 458, X.12
Eckert *vs*. Long Island Railroad: 352, VIII.1
economía del comportamiento: 33, I.4; 297, cap. VI, prob. 15; 580, XIV.4; 696, XVI.5; 727, XVI.15. *Véanse también* efecto de dotación; finanzas: del comportamiento
economía, metodología de la: 31, I.3
economías de escala y deseconomías de escala: 74, III.6; 86, III.8; 444, X.10; 455, X.11; 537, XIII.1; 867, XX.7; 928, XXII.12; 1057, XXVII.1
económica, renta, *véase* renta económica
educación: 209, V.4; 227, cap. V, prob. 2; 306, VII.2; 515, XII.8; 521, XII.10; financiamiento con impuestos locales a la propiedad: 749, XVII.5; como método de combate de la drogadicción: 339, VII.11; escuelas parroquiales: 1123, XXIX.8; 1088, segregada: XXVIII.2. *Véase también* inteligencia
efecto de dotación: 33, I.4; 114, III.14; 521, XII.9
eficiencia: 26, I.2; concepto de Kaldor-Hicks: 26, I.2; 50, II.4; concepto de Pareto: 26, I.2; 159, IV.10; 1141, XXXI.1; dependencia de, respecto de la distribución de la riqueza: 26, I.2; relación con la moral: 380, VIII.9. *Véase también* contratos: incumplimiento eficiente
ejecución hipotecaria (anticompetitiva): 444, X.10. *Véase también* arreglos de ataduras; fusiones: verticales
El joven Frankenstein (Brooks): XI.5, n 24
El origen de las especies (Darwin): XVI.15
elasticidad de la demanda: 391, IX.1; 395, IX.2; 400, IX.4; 409, cap. IX, prob. 1; 410, X.1; 430, X.6-444, X.10; como base de los "precios de Ramsey": 405, IX.5; 548, XIII.5; elasticidad cruzada: 432, X.7; 769, XVIII.1; de drogas adictivas: 339, VII.11; de mercado *vs*. de empresa: 432, X.7. *Véase también* tributación: análisis de la incidencia
elasticidad de la oferta: 306, VII.2; 317, VII.3; 339, VII.11; 430, X.6-436, X.8; de corto plazo *vs*. de largo plazo: 947, XXII.15
elección de la ley: 959, XXII.19
emociones: 15, I.1; 366, VIII.5; 378, VIII.8; 380, VIII.9; 986, XXIII.6; relación de, con la eficiencia: 380, VIII.9; en la teoría de juegos: 33, I.4
empleo: 39, I.5; compensación: 267, VI.11;

359, VIII.3; 509, XII.5; 515, XII.8; 521, XII.10; 623, XV.3; 652, XV.9; 658, XV.11-668, XV.14; 791, XVIII.8; de las mujeres: 199, V.2; 204, V.3; 521, XII.10; premios de riesgo para el, peligroso: 273, VI.12; 359, VIII.3; 797, XVIII.9; a voluntad: 120, IV.1; 509, XII.5; 521, XII.10; 1054, XXVI.3
empresa dominante: 455, X.11; 470, XI.2
entrada: 306, VII.2; 391, IX.1; 410, X.1; 421, X.2; 436, X.8-444, X.10; 537, XIII.1; 558, XIII.9; 850, XX.3; 867, XX.7; 877, XX.9; barreras para: 33, I.4; 395, IX.2-400, IX.4; 410, X.1; 444, X.10; 893, XXI.4; control de, por los reguladores del servicio público: 548, XIII.5; regulación de: 555, XIII.7
envidia: 740, XVII.2; 816, XIX.1
equidad: 47, II.2; 593, XIV.9; doctrina de: 1117, XXIX.5
equilibrio: 15, I.1; 33, I.4; competitivo: 15, I.1; parcial: 775, XVIII.3
ERISA (Ley de Seguridad de los Ingresos de Jubilación del Empleado): 531, XII.11; 704, XVI.9; efectos económicos de: 531, XII.11; y ESOP: 618, XV.2; otorgamiento incompleto de los beneficios de la pensión: 531, XII.11; planes de pensión de beneficio definido *vs*. los de contribución definida: 531, XII.11
error inocuo: 985, XXIII.5; en consideraciones ante el tribunal de apelaciones: 956, XXII.18
error legal: 306, VII.2; 380, VIII.9; 899, XXII.2; 918, XXII.9; 947, XXII.15; 956, XXII.18; 1027, XXIV.10; tipo I *vs*. tipo II: 979, XXIII.4. *Véase también* evidencia (pruebas); procedimiento
ESOP (Plan de Propiedad de Acciones de los Empleados): 618, XV.2; prob. de la subdiversificación en: 702, XVI.8
ESOP (Plan de Propiedad de Acciones de los Empleados): 618, XV.2; problema de la subdiversificación en, 702, XVI.8
especificidad de los activos: 783, XVIII.5. *Véase también* capital humano
estacionamiento ilegal: 897, XXII.1
Estatuto de Reclamación por Agravios contra Extranjeros: 1155, XXXI.5; 1159, cap. XXXI, prob. 2
estudios de eventos: 468, X.1
Evans *vs*. Newton: 826, XIX.3
EVRA Corp. *vs*. Swiss Bank Corp.: 170, IV.11
excedente del consumidor: 26, I.2; 179, IV.14; 396, IX.3; 409 cap. IX, prob. 3
externalidades: en los casos de cuasidelitos: 294, VI.17; causadas por la población: 217, V.6; congestionamiento: 86, III.8; 217, V.6; 479, XI.4; 481, XI.5; creadas por el traslado de los honorarios de abogados: 928, XXII.12; definición: 91, III.9; del discurso: 1105, XXIX.2; enfermedades de transmisión sexual como fuente de: 220, V.7; externalización del costo como técnica reguladora del derecho común: 264, VI.9; 359, VIII.3; 1105, XXIX.2; interestatales: 1060, XXVII.2-1076, XXVII.10; litigio como fuente de: 172, IV.12; pecuniarias *vs*. tecnológicas: 15, I.1; positivas de la educación: 769, XVIII.1; de red: 455, X.11; responsabilidad limitada como fuente potencial de: 623, XV.3; solución de la fusión (propiedad única): 86, III.8; de la votación: 1103, XXIX.1. *Véase también* problema del gorrón
extorsión: 86, III.8; 116, cap. III, prob. 6; 336, VII.10

fallo declaratorio: 900, XXII.3
falsificación: 301, VII.1
federalismo: aplicación del derecho penal federal *vs*. el derecho estatal: 1062, XXVII.3; beneficios del bienestar: 1076, XXVII.9; implicaciones de la aversión al riesgo: 1057, XXVII.1; poder del Congreso para regular el comercio: 1080, XXVII.11; y respuesta a emergencias: 1060, XXVII.2; teoría económica de: 1057, XXVII.1; transferencias de agua entre cuencas: 1075, XXVII.8
fideicomisos: de caridad: 826, XIX.3; 830, XIX.5; 831, XIX.6; deber de imparcialidad del administrador fiduciario: 732, cap. XVI, prob. 3; de dinastía: 816, XIX.1; contra el dispendio: 831, XIX.6; inversión de los fondos de: 650, XV.8; inversión social de: 708, XVI.10; obligación del administrador fiduciario de diversificar la cartera: 650, XV.8; 704, XVI.9; privados: 830, XIX.5; 831, XIX.6; regla del hombre prudente: 704, XVI.9. *Véase también* administradores fiduciarios
fiduciario: 152, IV.8; 674, XV.15. *Véase también* costos de agencia; agencia, ley de; fideicomiso
financiamiento del consumidor: 576, XIV.3
finanzas: apalancamiento: 633, XV.5; 689, XVI.2; 710, XVI.11; 720, XVI.14; arbitraje

del riesgo: 696, XVI.5; beta: 684, XVI.1; 693, XVI.4; 701, XVI.7; compra apalancada: 633, XV.5; del comportamiento: 696, XVI.5; contrato de honorarios contingentes como instrumento financiero: 928, XXII.12; corporativas: 623, XV.3; desintermediación: 664, XV.13; deuda asegurada *vs*. no asegurada: 628, XV.4; deuda *vs.* capital social y razón de deuda a capital social: 623, XV.3-633, XV.5; 684, XVI.1; 689, XVI.2; 701, XVI.7; diseño de la cartera: 684, XVI.1; efecto del impuesto a las ganancias de capital: 780, XVIII.4; 802, XVIII.10; fondos de índices: 704, XVI.9; 732, cap. XVI, prob. 1; fraude bancario *vs*. robo de banco: 1062, XXVII.3; e inversión social: 708, XVI.10; tesis del mercado eficiente: 693, XVI.4-699, XVI.6; 710, XVI.11; tesis de Modigliani-Miller: 689, XVI.2; trueque entre riesgos y rendimientos: 684, XVI.1. *Véanse también* banca; diversificación; fideicomisos; tasas de interés; transacciones aseguradas

Fórmula de Hand: 229, VI.1; 237, VI.4; 251, VI.6; 266, VI.10; 568, XIV.1; 886, XXI.3; 897, XXII.1; 977, XXIII.3; aplicación al castigo penal: 306, VII.2; 326, VII.5; aplicación en los cuasidelitos intencionales: 285, VI.15; en los casos de cateos y confiscaciones: 1128, XXX.1; contrapartida en el derecho de la prueba: 986, XXIII.6;

fraude: como base para la perforación del velo corporativo: 646, XV.7; constructivo: 253, VI.7; corporativo: 664, XV.13; en el mercado: 710, XVI.11; como razón para anular el matrimonio: 199, V.2; transferencia fraudulenta: 253, VI.7; en la venta de valores: 253, VI.7; 696, XVI.5; 710, XVI.11. *Véase también* declaraciones falsas; protección del consumidor; valores, regulación de los

Fuentes *vs*. Shevin: 1054, XXVI.3; 1156, cap. XXVI, prob. 1

fundaciones caritativas: 829, XIX.4

George, Henry: 780, XVIII.4

gobierno: contratos con: 190, cap. IV, prob. 11; corporaciones públicas: 680, cap. XV, prob. 1; empleo público: 509, XII.5; 1004, XXVI.3; financiamiento: 267, VI.11; incentivos de: 74, III.6; 548, XIII.5; como monopsonista: 460, X.13; organización óptima de: 1036, XXV.2; tierras públicas: 114, III.14. *Véase también* Constitución; federalismo; legislación; separación de poderes; *y agencias y programas gubernamentales específicos*

gorroneo: 378, VIII.8; 444, X.10; 501, XII.2; 628, XV.4; 751, XVII.6; 850, XX.3; 883, XXI.2; 1123, XXIX.8; en el contexto del mantenimiento del precio de reventa: 426, X.4; como justificación del derecho de las quiebras: 628, XV.4; 633, XV.5; como obstáculo para la sindicalización: 501, XII.2

Groves *vs.* John Wunder Co.: 159, IV.10

grupos de interés: 593, XIV.9; 769, XVIII.1; 810, XVIII.12; 850, XX.3; 855, XX.4; 857, XX.5; 867, XX.7; 1009, XXIV.4; 1020, XXIV.7; analogía con la cartelización: 850, XX.3; y la Constitución: 1033, XXV.1; 1040, XXV.3; 1054, XXVI.3

Guille *vs*. Swan: 245, VI.5

Hadley *vs*. Baxendale: 170, IV.11; 253, VI.7; 351, VIII.1; 882, XXI.1

hallazgo, derechos de: 59, III.2; 108, III.12

Hand, Learned: 229, VI.1; 1105, XXIX.2

Hayek, Friedrich A.: 49, II.3; 364, VIII.4; 378, VIII.8; 615, XV.1

herencia: efectos de incentivo: 816, XIX.1; fideicomiso de dinastía: 816, XIX.1; impuestos de defunciones y donaciones: 816, XIX.1; sucesión intestada: 823, XIX.2; legados para propósitos: 830, XIX.5; motivación de dar un legado: 786, XVIII.6; 816, XIX.1; movimiento para abolir la tributación sucesoria: 816, XIX.1; porción de la viuda: 833, XIX.7; problema del heredero asesino: 823, XIX.2; regla contra las disposiciones que aten la propiedad de forma perpetua: 816, XIX.1; 831, XIX.6; tributación de: 816, XIX.1. *Véase también* testamentos

Hicks, John R.: 400, IX.4

hijos: 209, V4; 215, V.5. *Véanse también* adopción; maternidad subrogada

hiperracionalidad: 33, I.4

Hobbes, Thomas: 120, IV.1; 1134, XXX.3

Holmes, Oliver Wendell: 100, III.11; 159, IV.10; 172, IV.12; 178, IV.13; 185, IV.16; 429, X.5; 1047, XXVI.1; 1103, XXIX.1; 1141, XXXI.1

homosexualidad: chantaje y: 330, VII.6; 999, XXIV.2; economía de: 220, V.7; homosexuales como víctimas de delitos de odio:

330, VII.6; matrimonio homosexual: 224, V.8
honor, su papel en el sistema de venganza: 366, VIII.5

igualdad: ante la ley: 375, VIII.7; costos sociales de la desigualdad del ingreso: 740, XVII.2. *Véase también* discriminación; riqueza: distribución de
impedimento colateral: 944, XXII.14; 968, cap. XXII, prob. XI. *Véase también res judicata*
impuesto al ingreso: ganancias adicionales: 543, XIII.3; 791, XVIII.8; 797, XVIII.9; corporativo: 628, XV.4; 668, XV.14; 783, XVIII.5; 802, XVIII.10; 1071, XXVII.5; 1083, cap. XXVII, prob. 3; crédito fiscal por el ingreso ganado: 515, XII.8; deducción de las caridades: 797, XVIII.9; deducción médica: 797, XVIII.9; deducciones: 769, XVIII.1; 780, XVIII.4; 791, XVIII.8; 797, XVIII.9; definición del ingreso para los fines de: 791, XVIII.8; efecto de ingreso *vs*. efecto de sustitución de: 786, XVIII.6; 791, XVIII.8; efecto de la progresividad sobre el pago del impuesto: 805, XVIII.11; efecto de las tasas marginales mayores: 805, XVIII.11; exclusión del ingreso no pecuniario de la producción doméstica: 797, XVIII.9; exclusión del pago de daños por lesión personal: 791, XVIII.8; exención del ingreso doméstico: 217, V.6; gravación de loterías y otros premios: 791, XVIII.8; *vs*. impuesto al consumo: 786, XVIII.6; como método de redistribución de la riqueza: 805, XVIII.11; negativo: 755, XVII.7; 766, cap. XVII, prob. 5; progresividad de: 805, XVIII.11; progresivo *vs*. proporcional: 805, XVIII.11
In re Baby M: 215, V.5
In re UAL Corp.: 702, XVI.8
inalienabilidad: 56, III.1
indemnización: de los honorarios legales de la parte ganadora: 928, XXII.12
índice de Herfindahl: 410, X.1
indiferencia deliberada: 285, VI.15
inflación: efecto sobre la regulación de los servicios públicos: 543, XIII.3; 558, XIII.9; e impuesto progresivo al ingreso: 805, XVIII.11; riesgo de: 531, XII.11
información: confusa: XXII.8; dar indicios de que se posee: 170, IV.11; obligación de revelar: 134, IV.4; 147, IV.7; 152, IV.8; 170, IV.11; planificación centralizada *vs.* mercados: 364, VIII.4. *Véase también* costos de la información; negociación de funcionarios internos
ingreso, utilidad marginal del: 26, I.2; 273, VI.12; 317, VII.3; 396, IX.3; 400, IX.4; 740, XVII.2; 766, cap. XVII, prob. 5; 786, XVIII.6; 805, XVIII.11; 816, XIX.1. *Véase también* impuesto al ingreso; riqueza
inmigración: 217, V.6; 346, VII.13; 359, VIII.3; 515, XII.8; 744, XVII.3; 1004, XXIV.3; 1076, XXVII.9
inmunidad: 296, VI.18; 359, VIII.3; 1131, XXX.2
innovación: 217, V.6; 351, VIII.1; 370, VIII.6; 400, IX.4; 470, XI.2; 479, XI.4; 558, XIII.9; 744, XVII.3; en empresas de la nueva economía: 455, X.11; teorías de: 378, VIII.8; tributación que induce a la innovación: 555, XIII.7. *Véase también* patentes; secretos comerciales
instituciones financieras, regulación de: 710, XVI.11-730, XVI.16
intención: evitación consciente como una forma de: 326, VII.5; significado económico de: 285, VI.15; 306, VII.2. *Véase también* delitos: premeditados
interés previo al juicio: 910, XXII.6. *Véase también* tasas de interés
interpretación de estatutos y constituciones: constitucional: 857, XX.5; 860, XX.6; 1033, XXV.1; estricta *vs*. construcción laxa: 886, XXI.3; uso de la historia legislativa después de la promulgación: 855, XX.4
interrogatorio: 973, XXIII.2; 988, XXIII.7
inversión, *véase* finanzas; fideicomisos: inversión de los fondos de
Islandia medieval: 366, VIII.5; 378, VIII.8

Jacob & Youngs, Inc. *vs*. Kent: 179, IV.14
Javins *vs.* First National Realty Corp.: 766, cap. XVII, prob. 8
Johnson Products: 573, XIV.2
Jolls, Christine: 50, II.4
jubilación temprana: 521, XII.10. *Véase también* ERISA
jueces: comparados con los jurados, 951, XXII.16; 977, XXIII.3; compensación de: 370, VIII.6; 860, XX.6; dobleteo o segundo empleo: 860, XX.6; elegidos *vs*. designados: 874, XX.8, 977, XXIII.3; función de

utilidad de: 860, XX.6; incentivos de, para seguir el precedente: 839, XX.1; 883, XXI.2; magistrados: 947, XXII.15; 956, XXII.18; 1131, XXX.2; normas que gobiernan a: 364, VIII.4; 882, XX.1-897, XX.2, cap. XX, prob. 1. *Véase también* tribunales
Junta de Aeronáutica Civil: 396, IX.3; 1016, XXIV.6
Junta Nacional de Relaciones Laborales: 501, XII.2; 1020, XXIV.7; estructura administrativa de: 1020, XXIV.7
jurado: 951, XXII.16-953, XXII.17; función política de: 1036, XXV.2; limitaciones cognoscitivas de: 986, XXIII.6; renuncia de, a participar en los juicios de instrucción: 953, XXII.17
jurisdicción: diversidad de ciudadanía: 1057, XXVII.1; personal: 1074, XXVII.7; requerimientos de cantidad mínima en la controversia: 947, XXII.15. *Véase también* capacidad para presentar demandas
jurisprudencia: 375, VIII.7; 380, VIII.9
justicia correctiva: 49, II.3; 285, VI.15; 375, VIII.7
justicia distributiva: 50, II.4; 737, XVII.1-755, XVII.10; teoría contractual (rawlsiana) de: 746, XVII.4; 746, XVII.4, n 16; 766, cap. XVII, prob. 9; 766, cap. XVII, prob. 10
justicia: 26, I.2; 47, II.2; 50, II.4; correctiva: 49, II.3; 285, VI.15; 375, VIII.7; distributiva: 50, II.4; 737, XVII.1-764, XVII.10; en la sala del tribunal: 839, XX.1; teoría contractual (rawlsiana) de: 746, XVII.4; 746, XVII.4n; 766, cap. XVII, prob. 9; 766, cap. XVII, prob. 10; 805, XVIII.11

Kaldor-Hicks, eficiencia de: 26, I.2; 50, II.4
Katz *vs.* United States: 1128, XXX.1, n 2
Keeton-O'Connell, plan de: 279, VI.14
Kelo *vs*. City of New London: 74, III.6, n 30
Kentucky Fried Chicken of California, Inc. *vs.* Los Angeles County Superior Court: cap. VIII, prob. 8
Kerr S. S. Co. *vs*. Radio Corp. of America: 855, XX.4, n 23
Keynes, John Maynard: 15, I.1, n 2; 39, I.5; 693, XVI.4; 727, XVI.15; 789, XVIII.7; *The General Theory of Employment, Interest and Money (Teoría general del empleo, el interés y el dinero)*: 39, I.5; 727, XVI.15; 789, XVIII.7
Kleist, Heinrich von: 1067, XXVII.4
Knight, Frank H.: 15, I.1, n 2; 45, II.1, n 6; 56, III.1, n 3
Kodak: 488, XI.7

Laffey *vs.* Northwest Airlines, Inc.: 897, cap. XXII, prob. 8
Laidlaw *vs*. Organ: 147, IV.7; 170, IV.11
lamentación: 15, I.1; 33, I.4
legados caritativos en: 830, XIX.5-831, XIX.6
legislación: 847, XX.2; 850, XX.3; ambigüedad en: 860, XX.6. *Véase también* Congreso; democracia; elección pública; interpretación; política
LePage's Inc. *vs.* 3M (Minnesota Mining and Manufacturing Co.): 444, X.10
Ley Clayton: 497, XII.1
Ley de Ajuste Agrícola: 506, XII.3
Ley de Aranceles: 1150, XXXI.3
Ley de Comunicaciones: 558, XIII.9
Ley de Cuidado de la Salud a Bajo Precio u Obamacare: 1080, XXVII.11
Ley de Derechos Civiles: 1131, XXX.2
Ley de Discriminación en el Empleo por la Edad: 521, XII.10
Ley de Dispositivos de Seguridad: 582, XIV.5, n 17
Ley de Educación para Individuos con Incapacidades: 764, XVII.9
Ley de Especies en Peligro: 116, cap. III, prob. 13 y 14
Ley de Extensión del Periodo de Copyright Sonny Bono: 481, XI.5
Ley de Fraudes: 190, cap. IV, prob. 10
Ley de Glass-Steagall: 716, XVI.13; 720, XVI.14
Ley de inventos de América Leahy-Smith: 495, cap. XI, prob. 1
Ley de Juicios Rápidos: 919, XXII.10
Ley de la Demanda: 15, I.1; 490, XI.8; 582, XIV.5
Ley de la Libre Elección del Empleado: 506, XII.3
Ley de la Veracidad en los Préstamos: 576, XIV.3
Ley de los Estadunidenses con Discapacidades: 521, XII.10
Ley de Paga Igual: 521, XII.10
Ley de Prevención del Abuso y Protección al Consumidor: 642, XV.6
Ley de Protección al Paciente y Cuidado Asequible: 1080, XXVII.11

Ley de Quiebras: 633, XV.5; XV.5n; 642, XV.6; 716, XVI.13
Ley de Reducción del Déficit: 565, cap. XIII, prob. 8
Ley de Reforma de la Libertad bajo Fianza de 1984: 919, XXII.10
Ley de Reforma de la Quiebra: 642, XV.6
Ley de Reforma de las Campañas Bipartidistas (McCain-Feingold): 867, XX.7; 1103, XXIX.1
Ley de Seguridad del Ingreso del Empleado Jubilado, *véase* ERISA
Ley de Seguridad y Salud en el Empleo: 515, XII.8
Ley de Telecomunicaciones: 558, XIII.9
Ley de Valores: 710, XVI.11
Ley del Aire Limpio: 584, XIV.6; 596, XIV.10
Ley del Homestead: 66, III.4; 488, XI.7
Ley Federal de Pensiones: 531, XII.11
Ley Federal de Reclamaciones de Cuasidelitos: 1062, XXVII.3
Ley Federal del Salario Mínimo: 515, XII.8; 1052, cap. XXVI, prob. 2
Ley Harter: 190, cap. IV, prob. 7
Ley Nacional de Recuperación Industrial: 506, XII.3; 1047, XXVI.1
Ley Nacional de Relaciones Laborales: 497, XII.1; determinación de la unidad de negociación: 501, XII.2; efecto sobre la contratación de trabajadores de remplazo durante la huelga: 501, XII.2; prohibición de contratación de "perro amarillo": 501, XII.2; representación exclusiva de la unidad negociadora: 501, XII.2
Ley Norris-LaGuardia: 497, XII.1
Ley Price-Anderson: 680, cap. XV, prob. 8
Ley Sarbanes-Oxley de 2002: 664, XV.13
Ley Sherman, 410, X.1; 429, X.5; 497, XII.1; 501, XII.2. *Véase también* derecho antimonopólico
Ley Staggers Rail: 548, XIII.5; 565, cap. XIII, prob. 5, prob. 6
Ley Taft-Hartley: 497, XII.1; 536, cap. XII, prob. 1
Ley Wagner: 497, XII.1; 506, XII.3. *Véase también* Ley Nacional de Relaciones Laborales
leyes *antidumping* y de cuotas compensatorias: 1150, XXXI.3
leyes de control de armas: 334, VII.9
leyes de ofensores múltiples: 317, VII.3
liberalismo: 50, II.4; 1119, XXIX.6
libertad de contrato: 1047, XXVI.1
libertad de expresión: 1103, XXIX.1; 1117, XXIX.5; para blogs: 1127, cap. XXIX, prob. 7; casos de camiones de sonido: 1105, XXIX.2; diferentes foros para: 1112, XXIX.3; discurso comercial: 1115, XXIX.4; 1119, XXIX.6; discurso político: 1103, XXIX.1; 1115, XXIX.4; análisis costo-beneficio de: 1115, XXIX.4; fórmula económica para la regulación de: 1105, XXIX.2; como legislación para grupos de interés: 1103, XXIX.1; de periódicos *vs*. radiodifusoras: 1117, XXIX.5; pornografía: 1105, XXIX.2; quema de la bandera como forma de: 1105, XXIX.2; regulación administrativa *vs*. judicial de: 1112, XXIX.3; restricciones previas: 1112, XXIX.3; subsidios de: 1105, XXIX.2; 1122, XXIX.7
libertad: 1047, XXVI.1; política *vs*. económica: 1047, XXVI.1; 1119, XXIX.6
libertades civiles: 737, XVII.1; 1042, XXV.4. *Véase también* libertad de expresión
licencia de "envoltura": 132, IV.3
licencias en blanco para el ejercicio de los derechos de reproducción: 490, XI.8
Lincoln, Abraham: 867, XX.7
litigación: arreglos de honorarios contingentes: 928, XXII.12; en casos de discriminación: 521, XII.10; 536, cap. XII, prob. 5; "costos" rembolsables *vs*. honorarios de abogados: 941, XXII.13; de las agencias públicas: 1004, XXIV.3; derechos civiles: 941, XXII.13; descubrimiento anterior al juicio: 902, XXII.4; 910, XXII.6; 928, XXII.12; 973, XXIII.2; determinantes de la tasa de: 928, XXII.12; dominio eminente: 74, III.6; efecto del traslado de los honorarios de abogados sobre la tasa de litigación: 928, XXII.12; 941, XXII.13; frecuencia de las victorias del demandante: 1004, XXIV.3; frívola: 941, *XXII.13; gastos en: 924, XXII.11; gastos de*predadores en: 924, XXII.11; honorarios por la presentación de los juicios: 881, cap. XX, prob. 2; 914, XXII.7; 947, XXII.15; de los indigentes: 947, XXII.15; juicios por perjuicios: 928, XXII.12; oferta de fallo: 941, XXII.13; traslado de los honorarios de abogados: 928, XXII.12. *Véase también* apelaciones; arreglo
litigio de derechos civiles: 941, XXII.13
Lloyd's: 190, cap. IV, prob. 14

Lochner *vs.* New York: 1047, XXVI.1
Lojack: 334, VII.9
Los diez mandamientos (DeMille): 1123, XXIX.8

MacKinnon, Catharine: 1105, XXIX.2
Madison, James: 1057, XXVII.1
Malthus, Thomas Robert: 217, V.6
mandatos, órdenes: 91, III.9; y mano de obra: 568, XIV.1; preliminares: 900, XXII.3; 956, XXII.18. *Véase también* cumplimiento específico
Manne, Henry: 45, II.1
marcas comerciales: 468, XI.1; 488, XI.7; banca de: 468, XI.1; de características funcionales: 488, XI.7; su efecto sobre la calidad del producto: 573, XIV.2
marginal, ingreso: 430, X.6; 458, X.12; definición de: 391, IX.1; y el cálculo del precio óptimo del monopolio: 395, IX.2; 775, XVIII.3; y el Teorema de Coase: 68, III.5
maternidad subrogada: 215, V.5
Mathews *vs*. Eldridge: 897, XXII.1
matrimonio: 227, cap. V, prob. 3; acuerdos premaritales de seguridad: 204, V.3; búsqueda premarital: 153, IV.9; 199, V.2; economía de: 194, V.1; homosexual: 224, V.8; naturaleza del contrato de: 199, V.2; polígamo: 220, V.7; 224, V.8; privilegios testimoniales: 990, XXIII.8; violación dentro de: 301, VII.1
McCarthy, Joseph: 1119, XXIX.6
mediación: 506, XII.3; 944, XXII.14; como una ayuda para el acuerdo: 916, XXII.8
Medicaid: 460, X.13; 764, XVII.10
Medicare: 460, X.13; 547, XIII.4; como contribuidor a la reducción de la tasa de pobreza: 751, XVII.6; como parte de la disyuntiva de los beneficios adquiridos: 764, XVII.10
mercados de capital, *véase* finanzas
mercados de trabajo: 497, XII.1; abuso del empleado: 521, XII.10; determinación del salario en: 521, XII.10; producción de equipo: 658, XV.11; salario de eficiencia: 521, XII.10. *Véase también* derecho del empleo
mercados de valores, *véase* corporaciones; finanzas; valores; valores, regulación de los
mercados negros: 209, V.4
mercantilismo: 15, I.1
MGM Studios, Inc. *vs*. Grokster, Ltd.: 481, XI.5
Michael Kohlhaas (Kleist): 468, XI.1; 1065, XXVII.4
Michelman, Frank I.: 56, III.1, n 2
Mickey Mouse: 479, XI.4; 481, XI.5
Microsoft: 488, XI.6; 668, XV.14
Mill, John Stuart: 1119, XXIX.6; *On Liberty*: 26, I.2; 224, V.8; 1047, XXVI.1; su opinión sobre los sindicatos: 497, XII.1; *Principles of Political Economy*: 766, cap. XVII, prob. 3
Miller, Henry: 1115, XXIX.4
Milton, John: 1112, XXIX.3
modernización, *véase* derecho
Modigliani-Miller, tesis de: 689, XVI.2
Mona Lisa: 835, cap. XIX, prob. 5
Monarco *vs.* Lo Greco: 823, XIX.2
monopolio bilateral: 74, III.6; 81, III.7; 91, III.9-100, III.11; 153, IV.9. *Véase también* monopolio
monopolio: 391, IX.1-407, IX.6; y durabilidad: 407, IX.6; bilateral: 74, III.6; 153, IV.9-170, IV.11; 199, V.2; 351, VIII.1; 501, XII.2; sobre la calidad y variedad del producto: 400, IX.4; compañías de derecho de vía: 74, III.6; competencia por: 405, IX.5; 455, X.11; daños contractuales del monopolista: 159, IV.10; del delito: 336, VII.10; determinación del precio monopólico que maximiza las ganancias: 391, IX.1; 410, X.1; 432, X.7; 1047, XXVI.1; discriminación de precios por el dueño de: 405, IX.5; sobre la discriminación racial: 1086, XXVIII.1; efecto de riqueza de: 396, IX.3; efecto sobre el contenido de las radiodifusiones: 1117, XXIX.5; efecto sobre el costo: 400, IX.4; sobre la entrada: 391, IX.1; de estándares: 455, X.11; fraude por el dueño de: 147, IV.7; ganancias descontadas en los precios de las acciones: 668, XV.14; gubernamental: 1057, XXVII.1; sobre la innovación: 400, IX.4; natural: 391, IX.1; 423, X.3; 455, X.11; 537, XIII.1; 557, XIII.8; 558, XIII.9; 1047, XXVI.1; pérdida de peso muerto causada por: 396, IX.3; precios depredadores como método de aseguramiento o protección de: 438, X.9; prob. del segundo mejor: 396, IX.3; producción bajo: 391, IX.1; religioso: 1123, XXIX.8; retardo de la entrada por: 444, X.10; cómo se define y se mide el poder monopólico: 430, X.6; de situación (oportunista): 120, IV.1; 153, IV.9; sucesivo: 444, X.10; teoría de: 391, IX.1-405, IX.6; tributación de: 775, XVIII.3.

Véanse también competencia: monopólica; discriminación de precios; monopsonio
monopsonio: 458, X.12; 1150, XXXI.3; en el mercado farmacéutico: 460, X.13; en los mercados laborales: 497, XII.1
monumentos, estatutos sobre: 74, III.6
moral, moralidad: 886, XXI.3; afectada por el número de religiones: 386, cap. VIII, prob. 5; 1123, XXIX.8; contratos implícitos: 184, IV.15; relación de, con la eficiencia: 380, VIII.9. *Véase también* utilitarismo
Morgan Stanley: 716, XVI.13
Morin Building Products Co. *vs*. Baystone Construction, Inc.: 120, IV.1
movimiento de código abierto (software): 488, XI.6
Muller *vs.* Oregon: 1047, XXVI.1
multas: 209, V.4; 301, VII.1-317, VII.3; 348, cap. VII, prob. 1; 370, VIII.6; 592, XIV.8; 979, XXIII.4; 996, XXIV.1
muros de partición: 94, III.10

National Collegiate Athletic Association *vs*. Board of Regents: 466, cap. X, prob. 1
negativa de responsabilidad en los casos de productos: 251, VI.6
negligencia: 194, VI.1; 237, VI.4; 279, VI.14; 331, VII.7; 351, VIII.1; asesinato negligente por el heredero: 823, XIX.2; componente de responsabilidad estricta en: 229, VI.1; en el derecho penal: 331, VII.7; evaluación del sistema: 279, VI.14; un prob. legal o fáctico: 543, XXIII.3; relevancia de los costos de la información: 234, VI.2; relevancia del riesgo de lesión para el victimario: 229, VI.1; *vs*. la responsabilidad estricta: 245, VI.5; 359, VIII.3; 918, XXII.9. *Véase también* Fórmula de Hand
negociación colectiva: 501, XII.2-507, XII.4; 514, XII.7; 536, cap. XII, prob. 1. *Véase también* sindicatos
negociación de funcionarios internos: 663, XV.12
negociación de la declaratoria de culpabilidad: 919, XXII.10
negociación de los que están adentro de la empresa: 663, XV.12
negociación exclusiva: 120, IV.1; 185, IV.16; 455, X.11
New State Ice Co. *vs*. Liebmann: 1047, XXVI.1
Newton, Isaac: 31, I.3
Nirvana y economía: 727, XVI.15
no lucrativas, *véase* corporaciones (empresas): no lucrativas
nómina: 501, XII.2; 515, XII.8; 521, XII.10; impuestos a: 764, XVII.10; 904, XXVII.5
normas sociales, *véase* normas
normas: cómo las afecta la competencia religiosa: 1123, XXIX.8; y derecho: 378, VIII.8; eficiencia de: 380, VIII.9; 615, XV.1; establecimiento de: 364, VIII.4. *Véase también* derecho común: como costumbre
Northern Securities Co. *vs*. United States: 429, X.5
nueva economía, probs. antimonopólicos en la: 455, X.11
Nuevo Trato o *New Deal*: 506, XII.3; 1020, XXIV.7; 860, XXVI.1

obscenidad: 1112, XXIX.3; 1119, XXIX.6; 1127, cap. XXIX, prob. 4
ocio: 267, VI.11; e impuesto al ingreso: 786, XVIII.6; 791, XVIII.8; como parte de la función de utilidad judicial: 860, XX.6
Odisea: 481, XI.5
Office Depot: 432, X.7
OfficeMax: 432, X.7
Oficina Administrativa de los Tribunales de los Estados Unidos: 979, XXIII.4, n 15
Oficina de Patentes y Marcas Registradas: 468, XI.1; 470, XI.2; 481, XI.5; 1025, XXIV.9
oligopolio: 421, X.2; teoría de la interdependencia de: 429, X.5
opiniones consultivas: 900, XXII.3
Orden Fraternal de las Águilas: 1123, XXIX.8
órdenes de detención, *véase* cateos y confiscaciones
órdenes: 91, III.9; preliminares: 900, XXII.3; 956, XXII.18. *Véase también* cumplimiento específico
Organización de Países Exportadores de Petróleo: 410, X.1
Organización del Tratado del Atlántico Norte: 466, cap. X, prob. 3
Organización Mundial del Comercio: 1144, XXXI.2
Organización Paternidad Planificada: 668, XV.14
organizaciones de derechos de reproducción: 490, XI.8
Organizaciones Regionales de Transmisión: 1065, XXVII.4

organizaciones: 680, cap. XV, prob. 1, prob. 6 y prob. 15; administración por excepción: 956, XXII.18; criminales: 336, VII.10; cultura de: 400, IX.4; economía de: 490, XI.8; 618, XV.2; gubernamentales *vs.* de corporaciones o empresas: 1036, XXV.2; judiciales *vs.* corporativas: 956, XXII.18; organización jerárquica de: 615, XV.1; 1057, XXVII.1. *Véase también* corporaciones (empresas)
ostracismo como aplicación de las normas: 370, VIII.6; 378, VIII.8

pago de indemnización por despido: 204, V.3; 652, XV.9
pagos de transferencia: 15, I.1; 237, VI.4; 285, VI.15; 366, VIII.5; 396, IX.3; 897, XXII.1
países en vías de desarrollo, *véase* derecho: y modernización
Paradine *vs*. Jane: 190, cap. IV, prob. 2
Pareto, concepto de: 26, I.2
parodia: 481, XI.5
Partido Comunista: 1105, XXIX.2; 1115, XXIX.4
Partido Demócrata: 830, XIX.5
Partido Republicano: 877, XX.9
patentes: 468, XI.1-478, XI.3; 481, XI.5; 490, XI.8; 1025, XXIV.9; bloqueo: 490, XI.8; combinaciones de: 490, XI.8; función económica de: 468, XI.1-458, XI.2; innovación simultánea: 478, XI.3; licenciamiento de: 490, XI.8; marañas de: 470, XI.2; con mejoras: 470, XI.2; obvias: 470, XI.2; reglas básicas de: 470, XI.2
paternalismo: 380, VIII.9; 576, XIV.3; 826, XIX.3; 831, XIX.6
Patrulla Fronteriza de los Estados Unidos: 1004, XXIV.3
Peabody: 100, III.11; 812, cap. XVIII, prob. 11
Peevyhouse *vs*. Garland Coal and Mining Co.: 159, IV.10
pena capital: 306, VII.2; 874, XX.8
pensión alimenticia, pago de: 204, V.3. *Véase también* divorcio
pérdida de peso muerto: 306, VII.2; 396, IX.3; 633, XV.5; 684, XVI.1
perdones: 325, VII.4
periódicos: aplicación de la doctrina de la equidad a: 1117, XXIX.5; 1127. cap. XXIX, prob. 6; y limitaciones a los remedios legales para la difamación: 1105, XXIX.2; responsabilidad por publicar nuevamente material difamatorio: 292, VI.16; subsidiados por la limitación fijada sobre la publicidad de campaña: 867, XX.7
perjuicio atrayente: 68, III.5
perjuicio: 81, III.7; *vs*. allanamiento: 91, III.9; atrayente: 68, III.5; 237, VI.4; basado en el temor irracional: 116, cap. III, prob. 9; 593, XIV.9; doctrina de "ir hacia…": 81, III.7; enfoque del equilibrio: 81, III.7; juicios de: 928, XXII.12; librería pornográfica como un: 1127, cap. XXIX, prob. 2; llegando a un arreglo respecto a: 81, III.7; manifestación pública de pornografía como un: 1112, XXIX.2
Pierson *vs*. Post: 59, 584, III.2
Pigou, A. C.: 45, II.1; 94, III.10; 584, XIV.6
plan de pensiones públicas: 531, XII.11; 746, XVII.4; 746, XVII.4, n 18
Plant, Arnold: 45, II.1
Ploof *vs*. Putnam: 237, VI.4; 264, VI.9
pobreza: 751, XVII.6-755, XVII.7; como causa del delito: 348, cap. VII, prob. 8; 751, XVII.6; pobres como una clase constitucionalmente protegida: 1052, XXVI.2; requerimiento de residencia para el subsidio a los pobres: 1076, XXVII.9. *Véase también* asesoría, derecho a la; bienestar; política de vivienda
poder de negociación: 153, IV.9; 386, cap. VIII, prob. 1; 766, cap. XVII, prob. 8; en el matrimonio: 194, V.1. *Véase también* monopolio bilateral
política de vivienda: aplicación del código de vivienda: 759, XVII.8; complementos de renta: 759, XVII.8; desalojo en venganza: 759, XVII.8; garantías de habitabilidad: 759, XVII.8; 766, cap. XVII, prob. 8; ordenanzas de derechos de los inquilinos: 759, XVII.8; ordeña: 759, XVII.8; vivienda pública: 755, XVII.7; 759, XVII.8
política poblacional: 217, V.6
política: barreras para la entrada en: 877, XX.9; competencia entre los partidos: 877, XX.9; enfoque antimonopólico para la regulación de: 877, XX.9; límites del mandato: 874, XX.8-877, XX.9; reforma del financiamiento de las campañas: 867, XX.7; teoría económica de la democracia: 867, XX.7. *Véase también* elección pública; legislación; libertad de expresión

Pomer *vs.* Schoolman: 386, cap. VIII, prob. 2
pornografía: 220, V.7; 1122, XXIX.7; regulación de: 1105, XXIX.2. *Véase también* obscenidad
Posadas de Puerto Rico Assocs. *vs*. Tourism Co. of Puerto Rico: 1115, XXIX.4; 1119, XXIX.6
posesión: 108, III.12; *vs*. títulos de papel como una base de los derechos de propiedad: 108, III.12. *Véase también* propiedad
Pound, Roscoe: 860, XX.6
Prah *vs.* Maretti: 116, cap. III, prob. 10
precedente: 47, II.2; 839, XX.1; 860, XX.6; 882, XXI.1-883, XXI.2; 944, XXII.14; distinguir *vs.* anular: 883, XXI.2. *Véase también stare decisis*
precio depredador: 410, X.9; análisis mediante la teoría de juegos de: 410, X.9; dumping como forma de: 1150, XXXI.3; prueba de Areeda-Turner: 410, X.9; pruebas legales de: 410, X.9; remedio de pago de daños por: 461, X.14. *Véase también* litigación: gastos depredadores en
precio: como contenedor de información: 74, III.6; 251, VI.6; 615, XV.1; en dos partes: 548, XIII.5; de equilibrio: 15, I.1; fijación del de servicios de red: 558, XIII.9; fijado de acuerdo con el costo marginal en condiciones de costo medio declinante: 548, XIII.5; fijado de acuerdo con el costo más un agregado: 421, X.2; 543, XIII.3; 678, XV.17; 962, XXII.20; fijado de acuerdo con el costo medio: 548, XIII.5; límite: 436, X.8; pleno *vs*. nominal: 339, VII.11; 488, XI.7; de Ramsey: 548, XIII.5; 555, XIII.7; regulación de precio tope: 547, XIII.4; relación del monopólico con el competitivo: 537, XIII.1; relativo: 15, I.1; sombra: 15, I.1; 94, III.10; de varias partes: 565, cap. XIII, prob. III. *Véase también* discriminación de precios; precio depredador; precios, controles de; regulación de tasas; competencia; monopolio
precios de Ramsey: 548, XIII.5
precios, controles de: 31, I.3; 209, V.4; análisis económico y efectos de: 15, I.1
premio de las acciones: 26, I.2
prescripción adquisitiva: 100, III.11
préstamos usurarios: 576, XIV.3; 227, cap. V, prob. 4
previsibilidad: 170, IV.11; 253, VI.7
Primera Enmienda: 50, II.4; 1040, XXV.3; 1105, XXIX.2; 1115, XXIX.4; 1117, XXIX.5; 1119, XXIX.6; 1123, XXIX.8; 1127, cap. XXIX, probs. 1, 3, 6, 7
principal no revelado: 152, IV.8
privacidad: 378, VIII.8; 479, XI.4; cateos y confiscaciones como una invasión de: 1128, XXX.1-1131, XXX.2; derecho de: 479, XI.4; 1128, XXX.1
privatización: 114, III.14; 558, XIII.9; 996, XXIV.1-999, XXIV.2; venta de caminos de peaje: 608, XIV.12
privilegios, *véase* evidencia
probabilidad: 15, I.1-26, I.2; 33, I.4; 137, IV.5; 194, VI.1; 237, VI.4; 979, XXIII.4; objetiva *vs.* subjetiva: 910, XXII.6; 928, XXII.12
problema del último periodo: 317, VII.3; 531, XII.11; 658, XV.11
ProCD, Inc. *vs.* Zeidenberg: 132, IV.3
procedimiento: administrativo: 1009, XXIV.4; 591, XXIV.7; adversarial *vs*. inquisitivo: 839, XX.1; 973, XXIII.2; descubrimiento previo al juicio: 902, XXII.4; 910, XXII.6; examen médico obligatorio del demandante: 910, XXII.6; impedimento colateral: 944, XXII.14; 968, cap. XXII, prob. 11; juicio por jurados: 951, XXII.16; juicio sumario: 902, XXII.4; juicios bifurcados: 924, XXII.11; metas económicas de: 897, XXII.1; moción de que se proceda a descartar: 902, XXII.4; obligación de hacer una indagación razonable: 968, cap. XXII, prob. 16; opiniones consultivas: 900, XXII.3; penal: 346, VII.13; 919, XXII.10; 979, XXIII.4; reglas del civil: 910, XXII.6; 941, XXII.13; 968, cap. XXII, prob. 16; réplicas obligatorias: 924, XXII.11; *res judicata*: 944, XXII.14. *Véase también* carga de la prueba; evidencia; jurado; jurisdicción
producción doméstica: teoría económica de: 194, V.1; 204, V.3; XII.10; tratamiento fiscal de: 791, XVIII.8; valuación de, en casos de cuasidelito: 267, VI.11; 297, cap. VI, prob. 8. *Véase también* derecho familiar y economía
productividad: del trabajo: 267, VI.11-273, VI.12; 497, XII.1; 507, XII.4; marginal: 497, XII.1; relacionada con la edad: 507, XII.4; 521, XII.10-531, XII.11
Producto Interno Bruto: 596, XIV.10; 608, XIV.12; 737, XVII.1, n 8
Producto Nacional Bruto: 658, XV.11
producto, diferenciación del: 410, X.1; 444, X.10; 547, XIII.4

productor, excedente del: 396, IX.3
productos, de marca *vs.* genéricos: 460, X.13; 573, XIV.2
profesión de abogado: 962, XXII.20; arreglos de honorarios contingentes: 285, VI.15; 928, XXII.12; asesoría de abogados sobre la evitación fiscal: 962, XXII.20; competencia dentro de: 962, XXII.20; conflictos de interés: 928, XXII.12; 962, XXII.20; costos de agencia: 916, XXII.8; 928, XXII.12; despachos de abogados de interés público: 968, cap. XXII, prob. 8; estructura económica de: 962, XXII.20; ética de: 962, XXII.20; papel de, en el crecimiento económico: 962, XXII.20; subsidio público de: 755, XVII.7; tamaño óptimo de: 962, XXII.20. *Véase también* despachos de abogados; servicios legales
prohibición del dopaje en el deporte profesional: 380, VIII.9
prohibición: 857, XX.5
Promedio Industrial Dow Jones: 710, XVI.11
promesa de impedimento colateral: 190, cap. IV, prob. 9; 366, VIII.5
promesas: 120, IV.1; 128, IV.2; 137, IV.5; 159, IV.10; 185, IV.16; 380, VIII.9; 1141, XXXI.1; 1144, XXXI.2; de autocumplimiento: 380, VIII.9; donativos: 823, XIX.2. *Véase también* contratos
propiedad comunitaria: 204, V.3
propiedad intelectual: 45, II.1; 56, III.1; 100, III.11; 108, III.12; 410, X.1; 455, X.11; 468, XI.1-490, XI.8, XXIX.2; doctrina de la apropiación indebida: cap. XI, prob. 6; doctrina del derecho moral: XI.1; en los programas de noticias: XXIX.2; en el software de computadora: 410, X.1; 455, X.11; 470, XI.2; 481, XI.5; 488, XI.6; 490, XI.8; XXIV.4. *Véase también* copyrights; marcas comerciales; patentes; secretos comerciales
propiedad: abandono de: 108, III.12; de aguas: 59, III.2; 66, III.4; 100, III.11; de animales: 59, III.2; asignación y definición de los derechos de: 15, I.1; 68, III.5; 113, III.13; de ballenas: 59, III.2; comunal: 94, III.10; construcción de la casa en terrenos públicos: 66, III.4; 488, XI.7; convenios restrictivos: 86, III.8; de cuevas: 116, cap. III, prob. 8; decisión de rentar o comprar: 791, XVIII.8; derecho de los desechos: 94, III.10; derecho de terrateniente-inquilino: 94, III.10; 59, III.12; derechos al uso futuro de: 66, III.4; derechos contractuales *vs*. derechos de: 86, III.8; 120, IV.1; derechos de *de facto vs. de jure*: 64, III.3; derechos de, *vs*. reglas de responsabilidad: 91, III.9; derechos sobre recursos naturales: 114, III.14; doctrina de la confianza pública: 59, III.2; doctrina de la servidumbre de vistas: 68, III.5; efectos de riqueza de: 113, III.13; del espectro electromagnético: 64, III.3; estatutos sobre monumentos: 74, III.6; funciones económicas de: 55, cap. III íntegro; heredades en la tierra (propiedad dividida): 94, III.10; historia de: 56, III.1; leyes del derecho de sembrar: 68, III.5; mejoras hechas por error: 116, cap. III, prob. 15; movimiento de código abierto: 488, XI.6; de obras de arte: 100, III.11; perdida *vs*. extraviada: 108, III.12; posesión como fuente de derechos de: 59, III.2; 108, III.12; registro de títulos: 86, III.8; regla de la captura: 59, III.2; significado de, en las cláusulas del debido proceso: 1054, XXVI.3; sistema de Torrens: 100, III.11; tenencia conjunta: 94, III.10; tenencia de por vida: 94, III.10; 816, XIX.1; tierras públicas: 114, III.14; transferencia de: 100, III.11; tributación: 780, XVIII.4; 805, XVIII.11; ventas interestatales de agua: 1075, XXVII.8. *Véase también* dominio eminente; posesión; propiedad intelectual; vivienda; *y doctrinas particulares del derecho de la propiedad*
Prosser, William L.: 229, VI.1
prostitución: 220, V.7; 336, VII.10; 1105, XXIX.2
protección ambiental: y análisis de opciones: 596, XIV.10; depósitos por recipientes obligatorios: 592, XIV.8; evaluación contingente: 593, XIV.9; papel del análisis costo-beneficio en: 593, XIV.9. *Véase también* calentamiento global; contaminación
protección del consumidor: 573, XIV.2
protección igual de las leyes: 375, VIII.7; 896, cap. XXI, prob. 4; 1040, XXV.3; 1044, XXV.5; 1056, cap. XXVI, prob. 2; 1091, XXVIII.3; comprensión económica de: 1112, XXIX.3
protección: 66, III.4; 159, IV.10; 769, XVIII.1
proteccionismo: 1150, XXXI.3
provocación y control de armas: 334, VII.9
prueba: carácter: 988, XXIII.7; delitos previos:

988, XXIII.7; económica en los casos antimonopólicos: 410, X.1-423, X.3; interrogatorio: 973, XXIII.2; 988, XXIII.7; 991, XXIII.9; modelos económicos de: 971, XXIII.1; opinión: 991, XXIII.9; privilegio de abogado-cliente: 990, XXIII.8; 962, XXII.20; privilegio contra la autoincriminación: 1134, XXX.3; privilegios maritales: 990, XXIII.8; probabilística: 979, XXIII.4; como proceso de búsqueda: 971, XXIII.1. *Véase también* carga de la prueba; error inocuo
pruebas, *véase* evidencia
publicidad, derecho de: 479, XI.4
publicidad: 438, X.9; 444, X.10; 565, cap. XIII, prob. 7; 573, XIV.2; 576, XIV.3; 850, XX.3; 867, XX.7; 1119, XXIX.6; 1122, XXIX.7; 1125, cap. XXIX, prob. 3; campaña: 867, XX.7; comercial: 850, XX.3; 1115, XXIX.4; falsa o engañosa: 147, IV.7; 573, XIV.2; 1119, XXIX.6; gastos de: 438, X.9; 444, X.10; 924. XXII.11; y libertad de expresión: 1105, XXIX.2; 1119, XXIX.6; oportunidades de empleo: 1093, XXVIII.4; regulación de casinos: 1115, XXIX.4

quiebra, bancarrota: 172, IV.12; 185, IV.16; 410, X.1; 623, XV.3; 628, XV.4; 720, XVI.14; del capítulo 11: 633, XV.5; del capítulo XIII 13: 642, XV.6; de Chrysler en 2009: 633, XV.5; cliente: 680, cap. XV, prob. 14; corporativa: 633, XV.5; Ley de Reforma de 1978: 642, XV.6; municipal: 680, cap. XV, prob. 12; personal: 642, XV.6; por qué la federal: 1062, XXVII.3; regla de la prioridad absoluta del derecho: 633, XV.5; regla del valor nuevo: 633, XV.5; *vs.* seguro: 633, XV.5; 692, XVI.3; de United Air Lines: 720, XVI.14
Quinta Enmienda, *véase* debido proceso; autoincriminación

racionalidad: restringida: 986, XXIII.6; en la ciencia económica: 15, I.1; 31, I.3; de la confesión de un delito: 1137, cap. XXX, prob. 1; el criminal como maximizador racional: 306, VII.2; de las dependencias aplicadoras de la ley: 576, XXIV.3; formal: 49, II.3; 375, VIII.7; de los precios depredadores: 410, X.9; revisión de: 1044, XXV.5; en la teoría de juegos: 33, I.4; de la venganza: 438, X.9
Raffles *vs*. Wichelhaus: 134, IV.4
Ragtime: 481, XI.5
Raintree y Starwood: 1156, XXXI.6
Rajam, Raghuram: 744, XVII.3
Rardin *vs*. T & D Machine Handling, Inc.: 351, VIII.1
Rawls, John: 375, VIII.7; 746, XVII.4
Reagan, Ronald: 47, II.2
recuperación: 153, IV.9; 179, IV.14; XXVI.3
recursos naturales: 100, III.11; 114, III.14; 1054, XXVII.4
Red Lion Broadcasting Co. *vs*. FCC: 1117, XXIX.5
Reed *vs.* King: 380, VIII.9
reforma del financiamiento de las campañas: 850, XX.3; 867, XX.7
regalías: 116, cap. III, prob. 12; 470, XI.2; 481, XI.5; 490, XI.8; 603, XIV.11; 812, cap. XVIII, prob. 11
regalos: 826, XIX.3; 829, XIX.4; 1091, XXVIII.3; de capital humano: 816, XIX.1; tributación de: 769, XVIII.1; 791, XVIII.8; 816, XIX.1; 835, cap. XIX, prob. 1, prob. 5, y prob. 6
Regina *vs*. Dudley and Stephens: 332, VII.8; 348, cap. VII, prob. 4
Regla 10b-5 de la Comisión de Bolsa y Valores: 710, XVI.11
regla contra perpetuidades: 100, III.11; 816, XIX.1; 831, XIX.6
regla de Durham: 326, VII.5, n 27
regla de la prioridad absoluta: 633, XV.5
regla de los beneficios colaterales: 278, VI.13; 279, VI.14
regla del compañero sirviente: 386, cap. VIII, prob. 2
regla en el caso de Shelley: 100, III.11
Reglas Federales de Evidencia: 971, XXIII.1; 985, XXIII.5; 986, XXIII.6; 988, XXIII.7; 991, XXIII.9
Reglas Federales del Procedimiento Civil: 910, XXII.6; 941, XXII.13; 1141, XXXI.1; 968, cap. XXII, probs. 1, 3, 16
reglas: 120, IV.1; 351, VIII.1; 359, VIII.3; 370, VIII.6; 568, XIV.1; 759, XVII.8; 847, XX.2; 860, XX.6; 914, XXII.7; 918, XXII.9; de avisos y comentarios: 1027, XXIV.10; *vs*. estándares: 370, VIII.6; 568, XIV.1; 783, XVIII.5; 839, XX.1; 886, XXI.3; 1042, XXV.4
regulación de los incentivos: 547, XIII.4
regulación de tasas: 543, XIII.3; determinación de las tasas por el valor del servicio: 548, XIII.5; estructuras de las tasas inducidas por la regulación: 548, XIII.5; implica-

ciones de la teoría de las finanzas para: 701, XVI.7; "llevarse lo mejor" : 555, XIII.7; precios de los servicios regulados: 540, XIII.2; prob. de la compensación justa: 540, XIII.2; regulación de los incentivos: 547, XIII.4; subsidio cruzado: 555, XIII.7; tasa de rendimiento: 540, XIII.2-543, XIII.3; traslado del costo de los combustibles: 547, XIII.4. *Véase también* precios de Ramsey
regulación del servicio público: 537, XIII.1-558, XIII.9; contribuciones caritativas: 812, cap. XVIII, prob. 9; control de abandono: 555, XIII.7; control de ganancias: 540, XIII.2-543, XIII.3; control de la estructura tarifaria y de entrada: 548, XIII.5; definición: 537, XIII.1; demora reguladora: 543, XIII.3-547, XIII.4; efecto Averch-Johnson: 543, XIII.3; regulación de razón deuda-capital accionario: 540, XIII.2
regulación: acreditación ocupacional: 568, XIV.1; de la banca: 576, XIV.3; 716, XVI.13; costos fijos de: 568, XIV.1; demanda de: 515, XIII.8; *vs*. el derecho común: 568, XIV.1; de los derechos de aterrizaje en los aeropuertos: 565, cap. XIII, prob. 4; del discurso: 1105, XXIX.2-1112, XXIX.3; efecto sobre el comportamiento de los administradores: 558, XIII.9; de la entrada en el negocio del hielo: 1047, XXVI.1; de los hospitales: 547, XIII.4; de medicamentos nuevos: 582, XIV.5; óptima: 568, XIV.1; de la pesca: 59, III.2; política de: 584, XIV.6; de los precios del gas natural: 74, III.6; seguridad y salud: 380, VIII.9; 497, XII.1; 515, XII.8; 582, XIV.5; *vs*. subsidios: 596, XIV.10; del uso del cinturón de seguridad: 582, XIV.5. *Véase también* desregulación; regulación del servicio público
reincidencia: 317, VII.3; 988, XXIII.7
reivindicación: 1054, XXVI.3
relaciones de principal-agente, *véase* agencia, costos de; agencia: derecho de; corporaciones
religión: creencias en el más allá: 172, IV.12; denominaciones o confesiones: 618, XV.2; 1123, XXIX.8; economía de la: 1123, XXIX.8; en escuelas privadas y públicas: 1123, XXIX.8; establecimientos de: 1123, XXIX.8; e instrucción escolar respecto al "diseño inteligente": 1123, XXIX.8; subsidios públicos: 1123, XXIX.8
remedios contractuales: 120, IV.1; 147, IV.7; 172, IV.12; 178, IV.13; 185, IV.16; 351, VIII.1; 380, VIII.9; 928, XXII.12; autoayuda: 179, IV.14; castigos y confiscaciones: 172, IV.12; cumplimiento específico: 178, IV.13; daños: 159, IV.10; daños consecuenciales: 170, IV.11; daños de esperanza *vs*. daños de confianza: 159, IV.10; 190, cap. IV, prob. 5; daños líquidos: 172, IV.12; daños punitivos: 172, IV.12; mitigación de daños: 159, IV.10; reformulación: 134, IV.4; *vs*. remedios de cuasidelitos: 351, VIII.1; restitución: 185, IV.16
remedios: 116, cap. III, prob. 15; 633, XV.5; 928, XXII.12; administrativos: 1020, XXIV.7; casos de cateos y confiscaciones: 1131, XXX.2; en los casos de discriminación: 1091, XXVIII.3-1093, XXVIII.4; del derecho común: 81, III.7; 568, XIV.1-573, XIV.2; equitativos: 178, IV.13; 461, X.14; interés previo al juicio: 910, XXII.6; mandatos preliminares: 900, XXII.3; 956, XXII.18; 968, cap. XXII, prob. 10; 1112, XXIX.3; publicidad correctiva: 573, XIV.2. *Véanse también* daños; mandatos; restitución
rendimientos decrecientes, ley de los: 615, XV.1
renta económica: 15, I.1; 481, XI.5; 521, XII.10; tributación de: 780, XVIII.4; 1065, XXVII.4. *Véase también* búsqueda de la renta
reorganización corporativa: 633, XV.5; 783, XVIII.5; falla de filtración: 633, XV.5
represalias: 366, VIII.5; 1144, XXXI.2; aranceles de: 1150, XXXI.3. *Véase también* venganza
representación exclusiva: 120, IV.1
reputación: 375, VIII.7; 509, XII.5; 991, XXIII.9; como garantía de credibilidad: 33, I.4; 410, X.9; de las naciones: 1144, XXXI.2. *Véase también* contratos: de autocumplimiento; difamación
res judicata: 944, XXII.14; 968, cap. XXII, prob. 14; y división de la reclamación: 944, XXII.14
rescate: 128, IV.2; 237, VI.4; 296, VI.18; 351, VIII.1; responsabilidad por no lograr hacerlo: 264, VI.9. *Véase también* hallazgo, derechos de; restitución; salvamento
Reserva Federal, *véase* Sistema de la Reserva Federal

responsabilidad estricta: actividades ultrapeligrosas: 245, VI.5; 568, XIV.1; analogía con el impuesto a la contaminación: 584, XIV.6; y causación: 245, VI.5; costos de litigación: 245, VI.5; por el daño causado por animales salvajes: 245, VI.5; en el derecho primitivo: 366, VIII.5; economía de, *vs*. la negligencia: 245, VI.5; efecto sobre el nivel de la actividad: 245, VI.5; efecto sobre el seguro: 245, VI.5; por lesiones causadas por las vacunas: 251, VI.6; *vs*. negligencia: 359, VIII.3; 839, XX.1; 918, XXII.9; responsabilidad en los productos: 251, VI.6; responsable superior como una forma de: 260, VI.8. *Véase también* contratos
responsabilidad limitada: 618, XV.2; 623, XV.3; perforación del velo corporativo: 646, XV.7
responsabilidad, *véase* responsabilidad limitada
responsable superior: 152, IV.8; 260, VI.8; 674, XV.15; 1131, XXX.2
restitución: 264, VI.9; 285, VI.15; en los casos contractuales: 159, IV.10; 185, IV.16; en los casos de *copyright*: 185, IV.16; en los casos penales: 317, VII.3; el error como razón para: 285, IV.15; ordenada por la Comisión Federal de Comercio: 573, XIV.2. *Véase también* remedios contractuales
retórica, economía de la: 973, XXIII.2
reventa, mantenimiento del precio de: 410, X.1; 426, X.4; efectos de bienestar de: 426, X.4; teoría de los servicios del distribuidor: 426, X.4; en toda la industria: 410, X.1
reversión al Estado: 59, III.2
Rhone *vs.* Stephens: 116, cap. III, prob. 12
Rickards *vs*. Sun Oil Co.: 253, VI.7
riesgo, aversión al: 26, I.2; 512, XII.6; de los abogados litigantes: 874, XXII.8; de los accionistas: 623, XV.3; de los administradores corporativos: 692, XVI.3; y arreglo de las demandas: 839, XX.1; 904, XXII.5; 874, XXII.8; y arreglos con honorarios contingentes: 928, XXII.12; y castigo penal: 306, VII.2; 886, XXI.3; y derecho de los cuasidelitos: 229, VI.1; efecto de, sobre el traslado de los honorarios legales: 928, XXII.12; efecto de, sobre la razón de precio-ganancias de las acciones: 689, XVI.2; implicaciones para el federalismo: 1057, XXVII.1; impuesto al ingreso: 802, XVIII.10; como incentivo para la diversificación: 684, XVI.1; como justificación de la responsabilidad limitada: 253, VI.7; papel de, en el diseño de la cartera: 684, XVI.1; 704, XVI.9; en la teoría de la justicia de Rawls: 746, XVII.4; 766, cap. XVII, probs. 9 y 10
riesgo, neutralidad al: 194, VI.1; 904, XXII.5
riesgo, preferencia por el: 26, I.2; 730, XVI.16; 737, XVII.1; de los delincuentes: 306, VII.2
riesgo: 15, I.1-26, 1.2; 74, III.6; 134, IV.4-142, IV.6, 159, IV.10; 194, VI.1; 237, VI.4; 267, VI.11-273, VI.12; 628, XV.4; 684, XVI.1; 737, XVII.1; actitud hacia: 137, IV.5; 267, VI.11; distinguido de la incertidumbre: 15, I.1; efecto del impuesto progresivo al ingreso sobre la asunción de riesgos: 805, XVIII.11; sistemático *vs*. diversificable: 623, XV.3; 684, XVI.1; 692, XVI.3; 704, XVI.9. *Véanse también* riesgo, aversión al; riesgo, neutralidad al; riesgo, preferencia por el
riesgos y primas salariales en el empleo peligroso: 737, XVII.1; 797, XVIII.9
riqueza, efectos de la: 33, I.4; 584, XIV.6; de los contratos: 120, IV.1; de la discriminación en reversa: 1093, XXVIII.4; 1096, XXVIII.5; del impuesto al ingreso: 555, XIII.7; de las leyes contra la discriminación: 1086, XXVIII.1; de las leyes de salario mínimo y jornada máxima: 847, XX.2; de la maternidad subrogada: 215, V.5; del monopolio: 396, IX.3; de la política de vivienda: 113, III.13; de las reglas de la responsabilidad: 759, XVII.8; de la "venta de bebés": 215, V.5. *Véase también* tributación
riqueza, maximización de la: 47, II.2; 380, VIII.9; 593, XIV.9; 766, cap. XVII, prob. 10; y violación: 301, VII.1. *Véase también* eficiencia: concepto de Kaldor-Hicks
riqueza: 15, I.1; 737, XVII.1; 749, XVII.5-751, XVII.6; 791, XVIII.8; 816, XIX.1-833, XIX.7; distribución de: 26, I.2; 108, III.13; 386, cap. VIII, prob. 3; 558, XIII.9; 746, XVII.4; 766, cap. XVII, prob. 3; falta de, como clave para entender el derecho penal: 301, VII.1-306, VII.2; y monto de las multas: 306, VII.2; 317, VII.3; redistribución de: 220, V.7; 306, VII.2; 642, XV.6; 740, XVII.2; 759, XVII.8; 789, XVIII.7; 847, XX.2; 1040, XXV.3. *Véanse también* justicia distributiva; pobreza
robo: comparado con el monopolio: 396, IX.3; costos sociales de: 285, VI.15; de obras de arte: 100, III.11

Rockefeller, John D.: 444, X.10
Roe *vs*. Wade: 220, V.7
ruido de aviones: 81, III.7
Ruperto del Rin, príncipe: 190, cap. IV, prob. 2
Ruxley Electronics & Construction Ltd. *vs*. Forsyth: 179, IV.14

salarios: 128, IV.2; 194, V.1; 267, VI.11; 359, VIII.3; 497, XII.1; 509, XII.5; 515, XII.8-521, XII.10; 558, XIII.9; 860, XX.6; 1047, XXVI.1; 1054, XXVI.3; 1086, XXVIII.1; 1093, XXVIII.4; determinantes de: 137, IV.5; 159, IV.10; 658, XV.11; significado económico de la "vacación pagada": 791, XVIII.8. *Véase también* capital humano; empleo
salvamento: 59, III.2; 153, IV.9; 351, VIII.1; 633, XV.5
Sandel, Michael: 215, V.5
Sanitary Farms Dairy, Inc.: 812, cap. XVIII, prob. 4
Schelling, Thomas C.: 347, cap. VII y 610, cap. XIV, *Lecturas recomendadas*
Schumpeter, Joseph A.: 400, IX.4; 455, X.11; 867, XX.7; 877, XX.9
secretos comerciales: 468, XI.1; 478, XI.3; 479, XI.4
segundo mejor, prob. del: 396, IX.3
seguridad social: 194, V.1; 224, V.8; 515, XII.8; 746, XVII.4; 764, XVII.10; 816, XIX.1; 1009, XXIV.4; beneficios para los discapacitados: 751, XVII.6; eliminación del tope al ingreso ganado: 791, XVIII.8
seguridad y salud, en el trabajo: 359, VIII.3; 515, XII.8. *Véase también* empleo
seguro: 68, III.5; 137, IV.5; antiseguro: 159, IV.10; seguro de responsabilidad: 279, VI.14; 692, XVI.3; autoaseguramiento: 137, IV.5; 253, VI.7; 797, XVIII.9; azar o riesgo moral y selección adversa: 137, IV.5; 142, IV.6; 751, XVII.6; cómo se determinan las primas: 279, VI.14; compra de, por las corporaciones: 137, IV.5; 692, XVI.3; criterios raciales en: 1100, cap. XXVIII, probs. 2 y 3; deducción del impuesto al ingreso de las pérdidas por siniestro, percance o desastre: 797, XVIII.9; contra el desempleo: 278, VI.13; doctrina de las desviaciones: 142, IV.6; para el dolor y el sufrimiento: 273, VI.12; exclusión del suicidio: 142, IV.6; límites o topes a las primas por mala práctica médica: 279, VI.14; litigación de la cobertura: 900, XXII.3; negligencia comparativa como una forma de: 237, VI.4; permisibilidad del descubrimiento acerca de la póliza de seguro de responsabilidad: 968, cap. XXII, prob. 3; Póliza del Bloqueo de los Joyeros: 142, IV.6; prueba genética: 142, IV.6; y regla de los beneficios colaterales: 278, VI.13; regulación de: 279, VI.14; responsabilidad: 279, VI.14; 692, XVI.3; responsabilidad estricta como una forma de: 245, VI.5; riesgo calculable: 142, IV.6; para la salud: 142, IV.6; 744, XVII.3; 791, XVIII.8; social: 746, XVII.4; 751, XVII.6; tarifa de la experiencia: 279, VI.14; 716, XVI.13; 1062, XXVII.3
selección adversa: 142, IV.6; 279, VI.14. *Véase también* protección del consumidor
Senado y senadores: 506, XII.3; 766, cap. XVII, prob. 2; 1033, XXV.1; 1036, XXV.2; 1033, cap. XXV, prob. 1
señalamiento: 172, IV.12; quema de la bandera como: 1005, XXIX.2
sentencia, véase delitos
separación de poderes: 1033, XXV.1; 1036, XXV.2; 1057, XXVII.1
Serrano *vs*. Priest: 1052, XXVI.2
Servicio de Impuestos Internos (IRS): 780, XVIII.4; 999, XXIV.2; 1009, XXIV.4; 1029, cap. XXIV, prob. 1; presupuesto: 769, XVIII.1; 999, XXIV.2; recompensa para informantes: 1029, cap. XXIV, prob. 2
Servicio Postal: 1062, XXVII.3
servicios legales, subsidio público de los: 755, XVII.7
servidumbre de vista": 68, III.5
servidumbres o permisos de ingreso: 56, III.1; 81, III.7; 86, III.8; 100, III.11
servidumbres: 56, III.1; 81, III.7; 86, III.8; 100, III.11
sesgo de confirmación: 33, I.4; 977, XXIII.3
sexo: ataques sexuales: 988, XXIII.7; regulación de: 220, V.7. *Véase también* aborto; derecho familiar y economía; homosexualidad; obscenidad; pornografía; violación
Shakespeare, William: 947, XXII.15; 1112, XXIX.3; *Hamlet*: 366, VIII.5; *Enrique V*: 370, VIII.6n; *El rey Lear*: 823, XIX.2; *Romeo y Julieta*: 481, XI.5
Shapiro *vs*. Thompson: 1076, XXVII.9
Sherwood *vs*. Walker: 134, IV.4
sida (y otras enfermedades de transmisión sexual): 220, V.7; 224, V.8

Simons, Henry C.: 45, II.1; 766, cap. XVII, *Lecturas recomendadas*
sindicatos: 507, XII.4; como cárteles laborales: 497, XII.1; discriminación racial de: 521, XII.10; 1086, XXVIII.1; 1091, XXVIII.3; economía de la organización de: 501, XII.2; economía de las huelgas: 497, XII.1, n 2; 501, XII.2; efecto de las leyes de salario mínimo sobre: 515, XII.8; efectos económicos de: 479, XII.1; efectos sobre la productividad: 507, XII.4; unión sindical: 501, XII.2; unión sindical cerrada: 536, cap. XII, prob. 1. *Véase también* derecho del empleo; Ley Nacional de Relaciones Laborales
Sistema de la Reserva Federal: 704, XVI.9; 720, XVI.14; 720, XVI.14, n 41; 744, XVII.3; 1009, XXIV.4; Junta de: 727, XVI.15; 789, XVIII.7
Smith, Adam: 15, I.1; 59, III.2; 351, VIII.1; 497, XII.1n. 6; 1123, XXIX.8; *An Inquiry into the Nature and Causes of the Wealth of Nations*: 15, I.1n. 6
Smith, señor A.: 646, XV.7
soborno: 370, VIII.6; 999, XXIV.2; aportaciones a una campaña como una forma de: 867, XX.7; corporativo: 674, XV.15
software de computadora: 455, X.11; 481, XI.5
Sony Corp. of America *vs.* University City Studios, Inc.: 481, XI.5
Sony: 481, XI.5
Stambovsky *vs.* Ackley: 380, VIII.9
Standard Fashion Co. *vs.* Magrane-Houston, Co.: 455, X.11
Standard Oil Trust: 444, X.10
Staples: 432, X.7
stare decisis: 839, XX.1; 893, XXI.4; 914, XXII.7; en Inglaterra: 928, XXII.12
State Farm Mutual Automobile Ins. Co. *vs.* Campbell: 285, VI.15
Stigler, George J.: 710, XVI.11; 1016, XXIV.6
subastas: 603, XIV.11; 928, XXII.12; de ondas aéreas: 64, III.3; uso de, para dar con el consejero de clase en las acciones de clase: 928, XXII.12
subsidio: 555, XIII.7; 744, XVII.3; 759, XVII.8; 810, XVIII.12; del discurso: 1122, XXIX.7; economía de, comparada con la regulación: 1122, XXIX.7; 766, cap. XVII, prob. 7; como justificación de las leyes *antidumping*: 1150, XXXI.3; por medio de las estructuras de la tasa regulatoria: 596, XIV.10; 612, cap. XIV, prob. 1; a la religión: 1123, XXIX.8; significado económico de: 359, VIII.3. *Véase también* bienestar
Sugar Institute, Inc. *vs.* United States: 466, cap. X, prob. 1
suicidio: 142, IV.6; 199, V.2; 380, VIII.9
superestrella, fenómeno de: 26, I.2
sustitutos: 15, I.1; 1065, XXVII.4; 1117, XXIX.5; 1150, XXXI.3. *Véase también* complementariedad
Sweatt *vs.* Painter: 1088, XXVIII.2

Tararí y Tarará, de *Alicia a través del espejo* (Carroll), replicados por los grandes partidos: 867, XX.7
tasas de interés, y regulación financiera: 710, XVI.11; componentes: 267, VI.11; determinantes de las tasas de interés pagadas por las corporaciones: 444, X.10; cómo las afecta la actitud hacia el riesgo, 267, VI.11; en bonos municipales: 805, XVIII.11; préstamos asegurados *vs.* Estatuto de Reclamación por Agravios contra Extranjeros: 1155, XXXI.5; 1159, cap. XXXI, prob. 2
tasas de interés: bonos municipales: 805, XVIII.11; cómo se afectan por la actitud hacia el riesgo: 267, VI.11; componentes de: 267, VI.11; determinantes de, pagadas por las corporaciones: 444, X.10; préstamos asegurados *vs*. no asegurados: 628, XV.4; y regulación financiera: 576, XIV.3
televisión: de paga: 554, XIII.6; regulación de, por cable: 603, XIV.11. *Véase también* transmisión
Teorema de Coase: 15, I.1; 33, I.4; 45, II.1; 68, III.5; 94, III.10; 159, IV.10; 410, X.1; 775, XVIII.3; 1057, XXVII.1
teoría de juegos: 33, I.4; 237, VI.4; análisis del chantaje por: 999, XXIV.2; aplicación al precio depredador: 410, X.9; aplicada al desacato: 968, cap. XXII, prob. 13; aplicada al otorgamiento del privilegio contra la autoincriminación forzosa: 1134, XXX.3; de la contratación del secretario legal: 512, XII.6; equilibrio de agrupación: 33, I.4; 1096, XXVIII.5; estrategias previas al compromiso: 410, X.9; inducción retroactiva: 33, I.4; juego de tribunal-agencia: 1027, XXIV.10; juego de tribunal-legislatura: 860, XX.6; 919, XXII.10
terrorismo: 317, VII.3; 344, VII.12; 1042, XXV.4; 1060, XXVII.2; castigo penal de: 344, VII.12

tesoros, caza de: 59, III.2; 116, cap. III, prob. 16; ocultos: 59, III.2
testamentos: 253, VI.7; 816, XIX.1-833, XIX.7; condiciones en: 826, XIX.3; 831, XIX.6; condiciones raciales en: 1091, XXVIII.3; condiciones religiosas en: 831, XIX.6; doctrina *Cy pres*: 826, XIX.3; 831, XIX.6; interpretación de: 823, XIX.2; legados caritativos en: 816, XIX.1; 826, XIX.3-829, XIX.4; prob. de la mano del muerto: 826, XIX.3; 831, XIX.6; rechazo del legado: 835, cap. XIX, prob. 4; restricciones a la enajenación: 831, XIX.6
testigos expertos, *véase* evidencia
transacciones aseguradas: 628, XV.4; reafirmación de la deuda: 680, cap. XV, prob. 14; regla de la prioridad absoluta: 628, XV.4; 633, XV.5; 732, cap. XVI, prob. 4
transferencias intergeneracionales, *véase* herencia
transmisión: 64, III.3; 554, XIII.6; regulación de: 115, XXIX.5. *Véase también* televisión
transporte colectivo: 45, II.1; 331, VII.7; 537, XIII.1; 548, XIII.5
trasplantes de órganos: 184, cap. IV, prob. 15
tratados: Naval de Washington: 1144, XXXI.2; y reclamación por agravios cometidos por extranjeros: 1155, XXXI.5; sanciones por incumplimiento de: 1144, XXXI.2
Tribunal de Apelaciones para el Circuito Federal: 1009, XXIV.4
tribunales: de apelación: 928, XXII.12; 956, XXII.18; cambios en la organización por la creciente carga de trabajo: 956, XXII.18; carga de trabajo: 860, XX.6; 947, XXII.15; 956, XXII.18; 1042, XXV.4; corrupción en: 370, VIII.6; demora en: 370, VIII.6; 910, XXII.6; 947, XXII.15; especializados: 956, XXII.18; 991, XXIII.9; en los Estados Unidos: 904, XXII.5; 979, XXIII.4; independencia de: 956, XXII.18; 857, XX.5; 860, XX.6; 874, XX.8; ingleses: 896, cap. XXI, prob. 2; 914, XXII.7; instrucción *vs*. apelación en la revisión de las decisiones administrativas: 956, XXII.18; mercado de contratación de los secretarios legales: 512, XII.6; opiniones judiciales: 47, II.2; 847, XX.2-850, XX.3; en los países en vías de desarrollo: 370, VIII.6; revisión abstracta por: 900, XXII.3. *Véase también* jueces
tributación: *ad valorem vs*. impuestos específicos planos: 775, XVIII.3; 812, cap. XVIII, prob. 10; al alcohol: 775, XVIII.3; 612, cap. XIV, prob. 14; análisis de la incidencia: 775, XVIII.3; 1065, XXVII.4; aplicación forzosa: 999, XXIV.2; 1004, XXIV.3; 1029, cap. XXIV, prob. 1; asesoría legal acerca de la evitación fiscal: 962, XXII.20; asignación entre los estados: 749, XVII.5; 1083, cap. XXVII, prob. 1 y 3; auditoría de los contribuyentes: 1004, XXIV.3; autoevaluación: 74, III.6; 780, XVIII.4; base fortificada o aumentada hasta el día de la muerte: 816, XIX.1; y caridad corporativa: 668, XV.14; a los cigarros: 775, XVIII.3; concesiones fiscales como atractivo para las empresas de otros estados: 1083, cap. XXVII, prob. 8; conscripción como forma de: 773, XVIII.2; de la contaminación: 584, XIV.6; de contribuciones a los planes de pensión: 531, XII.11; costo por preparar declaraciones de impuestos: 769, XVIII.1; costos administrativos: 769, XVIII.1; créditos fiscales para el cuidado de los hijos: 791, XVIII.8; criterio de un impuesto eficiente: 769, XVIII.1; cumplimiento riguroso de: 996, XXIV.1; cuota o peaje en carreteras: 663, XIV.12; de los daños causados por los cuasidelitos: 297, cap. VI, prob. 12; depósitos por recipientes obligatorios como una forma de: 592, XIV.8; discriminatoria: 746, XVII.4; doctrina de la sustancia sobre la forma: 664, XV.13; 783, XVIII.5; por el dominio eminente: 74, III.6; dominio eminente como una forma de: 74, III.6; 773, XVIII.2; efecto de, sobre la elasticidad cruzada de la demanda: 769, XVIII.1; efectos de riqueza sobre la eficiente: 769, XVIII.1; efectos de sustitución: 746, XXVII.4; de emisiones de dióxido de carbono: 596, XIV.10; a la energía eléctrica: 596, XIV.10; exención para las instituciones religiosas: 1123, XXIX.8; a la extracción: 1065, XXVII.4; 1073, XXVII.6; 766, cap. XXVII, prob. 1; fisuras y lagunas: 805, XVIII.11; ganancias de capital: 780, XVIII.4; 802, XVIII.10; 816, XIX.1; para impeler al logro de avances tecnológicos: 596, XIV.10; impuesto de capitación, per cápita: 769, XVIII.1; 773, XVIII.2; 816, XIX.1; 584, XIV.6; impuestos específicos: 339, VII.11; 775, XVIII.3; 812, cap. XVIII, prob. 6 y 10; 847, XX.2; 1065, XXVII.4; ingreso de renta imputado: 1086, XVIII.1; al ingreso ganado jugando Juegos

de Rol Online Masivos para Múltiples Reproductores (MMORPGs): 812, cap. XVIII, prob. 12; del ingreso no pecuniario de las familias: 267, VI.11; limitaciones de, como un método de redistribución de la riqueza: 847, XX.2; como método de redistribución de la riqueza: 847, XX.2; 850, XX.3; 867, XX.7; monopólica: 1065, XXVII.4; del monopolista: 775, XVIII.3; objetivos económicos de: 769, XVIII.1; patrimonio, herencia y donación: 816, XIX.1; predial: 1071, XXVII.5; probs. de transición: 786, XVIII.6; progresiva: 740, XVII.2; 805, XVIII.11; proporcional: 805, XVIII.11; por regulación: 555, XIII.7; 603, XIV.11; regulatoria (pigoviana) *vs*. de los ingresos: 584, XIV.6; de la renta económica: 775, XVIII.3; 780, XVIII.4; 1065, XXVII.4; de la riqueza: 1103, XIX.1; simplificación: 769, XVIII.1; como sustituto de la protección del *copyright* contra la grabación hogareña de videos: 612, cap. XIV, prob. 3; tarifa de desplazamiento: 591, XIV.7; tratamiento dispar de la deuda y el capital social: 628, XV.4; 689, XVI.2; tributación estatal del comercio interestatal: 1065, XXVII.4; al uso compensatorio: 1065, XXVII.4; de la venta o la reorganización corporativa: 783, XVIII.5; yendo en corto contra la caja: 769, XVIII.1. *Véase también* impuesto al ingreso

U.S. Golf Association *vs*. St. Andrews Systems, Data-Max, Inc.: 495, cap. XI, prob. 6
U.S. Steel: 421, X.2
Ulises (Joyce): 481, XI.5
Ultramares Corp. *vs*. Touche: 356, VIII.2
United Airlines: 702, XVI.8; 720, XVI.14
United Shoe Machinery Corporation: 444, X.10
United States *vs.* Arnold, Schwinn & Co.: 466, cap. X, prob. 1
United States *vs*. Booker: 919, XXII.10
United States *vs*. Carroll Towing Co.: 229, VI.1
United States *vs*. Container Corp. of America: 466, cap. X, prob. 1
United States *vs.* Continental Can Co.: 466, cap. X, prob. 1
United States *vs*. Correll: 797, XVIII.9, n 46
United States *vs*. Dennis: 1105, XXIX.2
United States *vs.* E. I. Du Pont de Nemours & Co.: 432, X.7
United States *vs*. General Electric Co.: 490, XI.8
United States *vs.* Lallemand: 330, VII.6, n 31
United States *vs*. Lopez: 1080, XXVII.11
United States *vs.* Reading Co.: 466, cap. X, prob. 1
United States *vs*. Sealy, Inc.: 466, cap. X, prob. 1
United States *vs.* Singer Mfg. Co.: 466, cap. X, prob. 1
United States *vs*. United Shoe Machinery Corp.: 444, X.10
United States *vs.* United States Gypsum Co.: 466, cap. X, prob. 1
United States *vs*. United States Steel Corp.: 421, X.2
Unocal: 1153, XXXI.4
uso justo (doctrina del *copyright*): 481, XI.5
usura: 642, XV.6; 227, cap. V, prob. 4
utilitarismo y utilidad: 15, I.1; 26, I.2; declinación en la utilidad marginal: 396, IX.3; definición: 26, I.2; del delincuente: 306, VII.2; efecto de, sobre la utilidad agregada de la desigualdad del ingreso: 740, XVII.2; esperada: 15, I.1; 746, XVII.4; 1004, XXIV.3; marginal decreciente: 26, I.2; 317, VII.3; 400, IX.4; 740, XVII.2; 766, cap. XVII, prob. 1; 805, XVIII.11; 816, XIX.1; utilidad media *vs.* total: 217, V.6; utilidades positivas y negativas interdependientes: 285, VI.15; 306, VII.2; 740, XVII.2; de la vida: 273, VI.12; de la violación para el violador: 301, VII.1

vaguedad: 421, X.2; 886, XXI.3
valor: 26, I.2; 317, VII.3; subjetivo: 74, III.6; 91, III.9; 179, IV.14; 301, VII.1; 773, XVIII.2; 839, XX.1. *Véase también* vida
valores, regulación de los: 710, XVI.11; daños *vs.* remedios de mandatos: 710, XVI.11; derecho común *vs*. modelo continental: 710, XVI.11; eficiente *vs.* ineficiente: 710, XVI.11; fraude en el mercado: 696, XVI.5; 710, XVI.11; 720, XVI.14; implicaciones de las finanzas conductuales para: 696, XVI.5; información considerada como sustancial: 710, XVI.11; de las nuevas emisiones: 710, XVI.11; requerimientos acerca del margen: 710, XVI.11; revelación obligatoria: 710, XVI.11; venta de todo el negocio: 710, XVI.11; de venta en corto: 696, XVI.5; 710, XVI.11; 720, XVI.14
valores: pagarés, obligaciones o bonos: 680,

cap. XV, prob. 11; 802, XVIII.10; analistas de: 664, XV.13; 710, XVI.11; diferencias de riesgo entre: 684, XVI.1. *Véase también* accionistas; corporaciones; valores, regulación de los
venganza: 410, X.9; economía de: 366, VIII.5; papel de la emoción en: 366, VIII.5; relación de, con la justicia correctiva: 883, XXXI.2
venta de derechos paternales: 215, V.5
ventas, derecho de las: 172, IV.12; aceptación de bienes defectuosos o poco satisfactorios por alguna otra razón: 179, IV.14; contratos de venta en abonos: 1054, XXVI.3; doctrina del comprador de buena fe: 59, III.11; doctrina del tenedor de buena fe: 153, IV.9
vida, valor de la: 273, VI.12; 580, XIV.4; 582, XIV.5; 612, cap. XIV, prob. 5
Vincent *vs*. Lake Erie Transp. Co.: 237, VI.4; 253, VI.7
violación: 301, VII.1; 988, XXIII.7; 1105, XXIX.2; estatutaria: 326, VII.5; 348, cap. VII, prob. 5; marital: 301, VII.1
votación: 652, XV.9; 850, XX.3; 1103, XXIX.1; extensión del sufragio: 1033, XXV.1; impuesto de capitación: 1052, XXVI.2; venta de votos: 646, XV.17

Wal-Mart: 444, X.10
Warren Earl: 346, VII.13. *Véase también* Corte Warren
Weber, Max: 49, II.3; 375, VIII.7; 1016, XXIV.6
West Coast Hotel Co. *vs.* Parrish: 1047, XXVI.1
West, Robin L.: 227, cap. V, prob. 7
Westinghouse: 490, XI.8
White *vs*. Jones: 356, VIII.2; 253, VI.7
Wickard *vs.* Filburn: 1080, XXVII.11
Williams *vs.* Illinois: 306, VII.2
Williamson *vs*. Lee Optical Co.: 1074, XXVI.1
Wood *vs*. Duff-Gordon: 120, IV.1
WorldCom: 664, XV.13

Ybarra *vs*. Spangard: 366, VIII.5

zonificación: 74, III.6; separación de usos *vs.* excluyente: 86, III.8

ÍNDICE GENERAL

Sumario 7
Prefacio 9

Primera Parte
EL DERECHO Y LA ECONOMÍA: INTRODUCCIÓN

I. *La naturaleza del razonamiento económico* 15
I.1. Conceptos fundamentales 15
I.2. Valor, utilidad, eficiencia 26
I.3. El realismo de los supuestos del economista 31
I.4. Irracionalidad e hiperracionalidad; el origen de la economía del comportamiento y de la teoría de juegos 33
I.5. Dos enfoques de la economía 39
Lecturas recomendadas 42
Problemas 43

II. *El enfoque económico del derecho* 45
II.1. Su historia 45
II.2. El análisis económico positivo y normativo del derecho . . 47
II.3. El enfoque continental 49
II.4. Críticas del enfoque económico del derecho 50
Lecturas recomendadas 52

Segunda Parte
EL DERECHO COMÚN

III. *La propiedad* 55
III.1. La teoría económica de los derechos de propiedad: aspectos estáticos y dinámicos 56
III.2. Problemas de la creación y la aplicación de los derechos de propiedad 59
III.3. Derechos de propiedad en el derecho y en la economía: frecuencias de transmisión y subastas de ondas de radio 64
III.4. Derechos al uso futuro 66
III.5. Usos incompatibles; una revisión del Teorema de Coase . . 68

III.6. El allanamiento y el dominio eminente 74
III.7. Perjuicio 81
III.8. Soluciones para el problema de los usos incompatibles de la tierra mediante convenios de fusión y restricción 86
III.9. Prohibiciones en comparación con el pago de daños como métodos de aplicación de los derechos de propiedad 91
III.10. La propiedad dividida: los derechos "divididos" sobre las tierras o bienes raíces 94
III.11. Problemas en la transferencia de los derechos de propiedad 100
III.12. La posesión 108
III.13. Los efectos distributivos de las asignaciones del derecho de propiedad 113
III.14. Las tierras públicas 114
Lecturas recomendadas 116
Problemas 116

IV. *Derechos y remedios contractuales* 120
IV.1. El proceso del intercambio y los papeles económicos del derecho contractual 120
IV.2. La consideración 128
IV.3. Problemas de formación; el origen del asentimiento mutuo y los contratos unilaterales 132
IV.4. El error mutuo 134
IV.5. Los contratos como seguros: las doctrinas de la imposibilidad y otras relacionadas 137
IV.6. Los contratos de seguros 142
IV.7. El fraude 147
IV.8. El derecho de la agencia 152
IV.9. Coacción, contratos impresos, poder de negociación e inconsciencia 153
IV.10. Principios fundamentales de los daños contractuales . . . 159
IV.11. Daños consecuenciales 170
IV.12. Castigos, daños punitivos, daños líquidos, confiscaciones y reglas de no modificación 172
IV.13. El cumplimiento específico 178
IV.14. Autoayuda, condiciones contractuales y cumplimiento sustancial 179
IV.15. Contratos implícitos 184
IV.16. La teoría de la opción del derecho contractual *vs*. la teoría moral 185

Lecturas recomendadas 189
Problemas 190

V. *El derecho familiar y el derecho sexual* 194
V.1. La teoría de la producción doméstica 194
V.2. Formación y disolución del matrimonio 199
V.3. Consecuencias de la disolución 204
V.4. La protección legal de los hijos 209
V.5. La maternidad subrogada 215
V.6. La ley y la población 217
V.7. ¿Por qué regular el comportamiento sexual? 220
V.8. El matrimonio homosexual 224
Lecturas recomendadas 226
Problemas 227

VI. *El derecho aplicable a los cuasidelitos* 229
VI.1. La economía de los accidentes y la Fórmula de Learned Hand de la responsabilidad por negligencia 229
VI.2. El estándar de la persona razonable 234
VI.3. La costumbre como una defensa 235
VI.4. Falta de la víctima: negligencia contribuyente y comparativa, asunción del riesgo y obligaciones para con los allanadores 237
VI.5. La responsabilidad estricta 245
VI.6. La responsabilidad por los productos 251
VI.7. Causación y previsibilidad 253
VI.8. Cuasidelitos conjuntos, contribución, indemnización; aplicación de la regla de que el superior responde y el hostigamiento sexual 260
VI.9. Rescate: responsabilidad contra restitución 264
VI.10. La función de las indemnizaciones por cuasidelitos . . . 266
VI.11. Indemnizaciones por pérdida de la capacidad de ganar ingresos 267
VI.12. Indemnizaciones por el dolor y el sufrimiento, el problema de la valuación de la vida humana y el riesgo de la compensación excesiva 273
VI.13. La regla de los beneficios colaterales (fuente colateral) 278
VI.14. Negligencia con seguro de responsabilidad y seguro contra accidentes, compensación por accidente automovilístico sin falta y limitación de las demandas por mala práctica médica 279
VI.15. Los cuasidelitos intencionales 285

VI.16. La difamación 292
VI.17. La responsabilidad del sucesor 294
VI.18. Exenciones 296
Lecturas recomendadas 297
Problemas 297

VII. *El derecho penal* 301
VII.1. La naturaleza y la función económicas del derecho penal 301
VII.2. Sanciones penales óptimas 306
VII.3. Prevención del delito: leyes de infractores múltiples, tentativa y conspiración, instigación y complicidad, reincidencia, entrampamiento 317
VII.4. Los perdones 325
VII.5. La intención criminal 326
VII.6. Los delitos de odio 330
VII.7. Descuido, negligencia y responsabilidad estricta 331
VII.8. La defensa de la necesidad (compulsión) 332
VII.9. Precauciones de las víctimas; acerca de la provocación y el control de armas 334
VII.10. La economía del crimen organizado 336
VII.11. La "guerra contra las drogas" 339
VII.12. El castigo penal de los terroristas 344
VII.13. Las olas delictivas 346
Lecturas recomendadas 347
Problemas 348

VIII. *El derecho común, la historia del derecho y la jurisprudencia* . . 351
VIII.1. La lógica económica implícita del derecho común . . . 351
VIII.2. Los cuasidelitos económicos 356
VIII.3. El derecho común, el crecimiento económico y la historia del derecho 359
VIII.4. El derecho común como costumbre: el cuestionamiento de Hayek del análisis económico del derecho común . . 364
VIII.5. El derecho primitivo 366
VIII.6. El derecho y la modernización 370
VIII.7. El imperio de la ley 375
VIII.8. El derecho y las normas sociales 378
VIII.9. El contenido moral del derecho común 380
Lecturas recomendadas 384
Problemas 386

Tercera Parte
REGULACIÓN PÚBLICA DEL MERCADO

IX. *La teoría del monopolio* 391
IX.1. El precio y la producción del monopolista 391
IX.2. Efecto de los cambios del costo o de la demanda sobre el precio de monopolio 395
IX.3. Consecuencias de eficiencia del monopolio; la pérdida de peso muerto y la búsqueda de renta 396
IX.4. Otras objeciones económicas contra el monopolio; origen de la innovación y la minimización del costo 400
IX.5. Discriminación de precios 405
IX.6. Dos impedimentos a la monopolización (además de la entrada): la competencia por el mercado y la durabilidad 407
Lecturas recomendadas 409
Problemas 409

X. *Las leyes antimonopólicas* 410
X.1. Los cárteles y la Ley Sherman 410
X.2. Colusión tácita 421
X.3. Problemas con las pruebas económicas 423
X.4. Mantenimiento del precio de reventa 426
X.5. Fusiones que crean monopolios u oligopolios 429
X.6. El poder de mercado o monopólico 430
X.7. Definición del mercado y participaciones en el en el mercado 432
X.8. La competencia potencial 436
X.9. La depredación 438
X.10. Ejecución hipotecaria, ataduras y agrupamientos, barreras a la entrada 444
X.11. El antimonopolio y la "nueva economía" 455
X.12. Boicots; origen del monopsonio 458
X.13. ¿Existe el monopsonio benigno? 460
X.14. Los daños antimonopólicos 461
Lecturas recomendadas 465
Problemas 466

XI. *La propiedad intelectual* 468
XI.1. ¿Las patentes son una propiedad? 468
XI.2. Patentes e innovación 470
XI.3. El secreto comercial 478

XI.4. La privacidad 479
XI.5. El *copyright* 481
XI.6. El movimiento de código abierto 488
XI.7. Las marcas comerciales 488
XI.8. Acuerdos de patentes, las licencias globales de BMI-ASCAP y la intersección entre el antimonopolio y la propiedad intelectual 490
Lecturas recomendadas 495
Problemas 495

XII. *La regulación de la relación del empleo* 497
XII.1. El tratamiento especial de los cárteles laborales 497
XII.2. La lógica económica de la Ley Nacional de Relaciones Laborales 501
XII.3. La Ley de la Libre Elección del Empleado 506
XII.4. Los sindicatos y la productividad 507
XII.5. Empleo a voluntad 509
XII.6. La falla del mercado laboral: el caso de los secretarios legales judiciales 512
XII.7. Los trabajadores y el derecho antimonopólico 514
XII.8. La legislación del salario mínimo y de la "protección del trabajador" 515
XII.9. Beneficios ordenados 521
XII.10. Discriminación en el empleo por la raza, el género, la edad y la discapacidad 521
XII.11. La Ley Federal de Pensiones 531
Lecturas recomendadas 535
Problemas 536

XIII. *El servicio público y la regulación del medio de transporte colectivo* 537
XIII.1. El monopolio natural 537
XIII.2. El control de las ganancias y el problema del rendimiento razonable 540
XIII.3. Otros problemas causados por el intento de limitar las ganancias de una empresa regulada 543
XIII.4. La regulación de los incentivos 547
XIII.5. Regulación de la estructura tarifaria y de la entrada . . 548
XIII.6. La televisión de paga 554
XIII.7. Tributación por regulación (subsidio interno o subsidio cruzado) 555
XIII.8. La demanda de regulación 557

XIII.9. El movimiento de desregulación; la privatización . . . 558
Lecturas recomendadas 565
Problemas 565

XIV. *La elección entre la regulación y el derecho común* 568
XIV.1. La regulación óptima 568
XIV.2. Una revisión del fraude al consumidor 573
XIV.3. La revelación obligatoria 576
XIV.4. Alternativas a la revelación: el caso de las grasas trans . 580
XIV.5. Seguridad y salud 582
XIV.6. Regulación de la contaminación 584
XIV.7. El congestionamiento del tránsito 591
XIV.8. Depósitos por recipientes obligatorios 592
XIV.9. El análisis costo-beneficio como una herramienta reguladora 593
XIV.10. Los límites del análisis costo-beneficio: el caso del calentamiento global 596
XIV.11. Televisión por cable: problemas de *copyright* y del monopolio local 603
XIV.12. La venta de las autopistas 608
Lecturas recomendadas 610
Problemas 612

Cuarta Parte
EL DERECHO DE LAS ORGANIZACIONES MERCANTILES Y LOS MERCADOS FINANCIEROS

XV. *Corporaciones, financiamiento asegurado y no asegurado, quiebra* 615
XV.1. La naturaleza de la empresa 615
XV.2. Las diferentes clases de empresas 618
XV.3. La corporación como un contrato estándar 623
XV.4. Deuda corporativa y deuda con respaldo específico . . 628
XV.5. La quiebra corporativa 633
XV.6. La quiebra personal 642
XV.7. Tras el velo corporativo 646
XV.8. La separación de la propiedad y el control 650
XV.9. La transferencia del control corporativo 652
XV.10. Expulsiones corporativas y la competencia en la laxitud 656
XV.11. El problema de la compensación 658

XV.12. La negociación de las personas que poseen información confidencial de la empresa y el problema de la recompensa empresarial 663
XV.13. Los escándalos corporativos 664
XV.14. La discreción gerencial y las responsabilidades sociales de la corporación 668
XV.15. La criminalidad corporativa 674
XV.16. La corporación cerrada 676
XV.17. Una revisión de la regulación del servicio público . . . 678
Lecturas recomendadas 679
Problemas 680

XVI. *Los mercados financieros* 684
XVI.1. El diseño de la cartera 684
XVI.2. Diversificación, apalancamiento y la razón de deuda-capital accionario 689
XVI.3. ¿Por qué compran seguros las corporaciones? 692
XVI.4. La selección de acciones y la tesis del mercado eficiente 693
XVI.5. El desafío de las finanzas del comportamiento 696
XVI.6. Un reto mayor 699
XVI.7. Implicaciones de la teoría financiera para el monopolio regulado 701
XVI.8. Una revisión del ESOP: el problema de la subdiversificación 702
XVI.9. La Ley de la Inversión en Fideicomisos y los fondos de índices 704
XVI.10. Inversión social por los administradores fiduciarios . . 708
XVI.11. La regulación de los mercados de valores 710
XVI.12. Una revisión de la compensación: el caso de las comisiones de los fondos mutuos 715
XVI.13. ¿Por qué regular a los bancos? 716
XVI.14. El desplome de 2008 720
XVI.15. La Falacia del Nirvana y la economía darwiniana . . . 727
XVI.16. ¿Son codiciosos los banqueros? 730
Lecturas recomendadas 731
Problemas 732

Quinta Parte

EL DERECHO Y LA DISTRIBUCIÓN DEL INGRESO Y LA RIQUEZA

XVII. *Las desigualdades del ingreso, la justicia distributiva y la pobreza* 737
XVII.1. La medición de la desigualdad 737
XVII.2. ¿La desigualdad es ineficiente? 740
XVII.3. El aumento de la desigualdad: causas y consecuencias 744
XVII.4. La teoría contractual de la justicia distributiva y su aplicación a la seguridad social 746
XVII.5. La movilidad social 749
XVII.6. Los costos de la pobreza y las limitaciones de la caridad privada 751
XVII.7. Transferencias ilimitadas en efectivo *versus* beneficios en especie 755
XVII.8. Redistribución de la riqueza por reglas de responsabilidad: el caso del cumplimiento del código de vivienda 759
XVII.9. Beneficios abiertos en especie 764
XVII.10. La disyuntiva de los beneficios por derecho 764
Lecturas recomendadas 766
Problemas 766

XVIII. *La tributación* 769
XVIII.1. Tributación y eficiencia 769
XVIII.2. La conscripción 773
XVIII.3. Impuestos específicos 775
XVIII.4. Impuestos prediales 780
XVIII.5. La tributación del ingreso corporativo 783
XVIII.6. Impuestos sobre la renta contra impuestos al consumo 786
XVIII.7. Impuesto al valor agregado 789
XVIII.8. La definición del ingreso 791
XVIII.9. Deducciones del impuesto al ingreso 797
XVIII.10. El tratamiento especial de las ganancias de capital . . 802
XVIII.11. El principio del impuesto progresivo 805
XVIII.12. Los subsidios fiscales 810
Lecturas recomendadas 811
Problemas 812

XIX. *La transmisión de la riqueza tras la muerte* 816
XIX.1. Impuestos a la sucesión (y las donaciones) 816
XIX.2. Herederos interesados: el caso del heredero asesino . . . 823
XIX.3. La mano del muerto y la doctrina *Cy pres* 826
XIX.4. Los incentivos de las fundaciones caritativas 829
XIX.5. Legados para propósitos 830
XIX.6. Condiciones de los fideicomisos privados: origen del fideicomiso contra el dispendio 831
XIX.7. La porción de la viuda 833
Lecturas recomendadas 834
Problemas 835

Sexta Parte
EL PROCESO LEGAL

XX. *El mercado, el sistema acusatorio y el proceso legislativo como métodos de la asignación de recursos* 839
XX.1. Comparación de la asignación legal y la del mercado; comparación de los sistemas jurídicos inquisitivo y acusatorio 839
XX.2. Comparación de la asignación judicial y la legislativa . . 847
XX.3. La teoría económica de la legislación 850
XX.4. La interpretación estatutaria en un mundo de política de grupos de interés 855
XX.5. El poder judicial independiente y la política de los grugrupos de interés 857
XX.6. ¿Qué maximizan los jueces? 860
XX.7. La teoría económica de la democracia 867
XX.8. Democracia, jueces electos y las limitantes del mandato judicial 874
XX.9. Un enfoque antimonopólico de la democracia 877
Lecturas recomendadas 880
Problemas 881

XXI. *El proceso de elaboración de las normas jurídicas* 882
XXI.1. El cuerpo de precedentes como un acervo de capital . . 882
XXI.2. La producción de precedentes 883
XXI.3. Producción estatutaria; normas contra estándares . . . 886
XXI.4. *Stare decisis* 893
Lecturas recomendadas 895
Problemas 896

XXII. *Procedimiento civil y penal* 897
XXII.1. Las metas económicas del procedimiento 897
XXII.2. Los costos del error en los casos civiles 899
XXII.3. La adjudicación anticipatoria. Origen de las órdenes preliminares, los fallos declaratorios y las opiniones consultivas 900
XXII.4. El juicio sumario y las peticiones de desestimación . . 902
XXII.5. La decisión de transigir o ir a juicio 904
XXII.6. Cómo afectan al arreglo las reglas del procedimiento 910
XXII.7. La decisión de llegar a un arreglo y la evolución del derecho común 914
XXII.8. La mediación como un auxiliar del arreglo 916
XXII.9. Una revisión de las reglas de la responsabilidad . . . 918
XXII.10. Negociación de la declaración de culpabilidad, la reforma del procedimiento penal y la tesis de la inutilidad 919
XXII.11. Los gastos en litigación y la búsqueda del procedimiento eficiente 924
XXII.12. El acceso a los remedios legales: honorarios contingentes, demandas colectivas e indemnización de los honorarios legales 928
XXII.13. La Regla 68 y la indemnización unilateral 941
XXII.14. La *res judicata* y el impedimento colateral 944
XXII.15. La demora judicial y las crisis de la carga de trabajo . 947
XXII.16. Jurados y árbitros 951
XXII.17. La adjudicación óptima: el caso del contrato 953
XXII.18. Las apelaciones 956
XXII.19. La elección de la ley 959
XXII.20. La economía de la profesión de abogado 962
Lecturas recomendadas 966
Problemas 968

XXIII. *La prueba* 971
XXIII.1. La prueba como búsqueda y como minimización del error 971
XXIII.2. Comparación ulterior de los modelos acusatorio e inquisitivo 973
XXIII.3. Una revisión del juicio por jurados 977
XXIII.4. La carga de la prueba 979
XXIII.5. El error inocuo 985
XXIII.6. La prueba del balanceo de la Regla 403 986
XXIII.7. La prueba del carácter 988

XXIII.8. El privilegio 990
XXIII.9. Los testigos expertos 991
Lecturas recomendadas 993
Problemas 994

XXIV. *La aplicación de la ley y el proceso administrativo* 996
XXIV.1. Aplicación pública de la ley contra aplicación privada: los dilemas 996
XXIV.2. El cumplimiento público *versus* el cumplimiento privado: implicaciones positivas 999
XXIV.3. Elección de casos por parte de las agencias públicas 1004
XXIV.4. La regulación de los tribunales *versus* la regulación de las agencias administrativas 1009
XXIV.5. La superposición de las agencias 1014
XXIV.6. La captura del regulador 1016
XXIV.7. La estructura de la agencia administrativa 1020
XXIV.8. El comportamiento de las agencias administrativas . 1023
XXIV.9. Los costos de agencia en las agencias públicas 1025
XXIV.10. La revisión judicial de las decisiones de las agencias . 1027
Lecturas recomendadas 1028
Problemas 1029

Séptima Parte
LA CONSTITUCIÓN Y EL SISTEMA FEDERAL

XXV. *La naturaleza y las funciones de la Constitución* 1033
XXV.1. La economía del constitucionalismo: una introducción 1033
XXV.2. La separación de poderes 1036
XXV.3. La protección de los derechos 1040
XXV.4. El alcance de los derechos constitucionales: el caso del terrorismo 1042
XXV.5. Revisión de la racionalidad 1044
Lecturas recomendadas 1045
Problema 1046

XXVI. *El debido proceso económico* 1047
XXVI.1. La libertad contractual como un principio constitucional 1047
XXVI.2. El resurgimiento del debido proceso económico: los pobres como una clase constitucionalmente protegida 1052

XXVI.3. Los derechos del debido proceso de los consumidores y los empleados públicos 1054
Lecturas recomendadas 1056
Problemas 1056

XXVII. *La economía del federalismo* 1057
XXVII.1. La asignación de responsabilidades entre el gobierno federal y los estados 1057
XXVII.2. La respuesta a las emergencias: Katrina (o Sandy) y el terrorismo 1060
XXVII.3. Los tribunales federales y la aplicación de la ley federal 1062
XXVII.4. Tributación estatal: impuestos específicos 1065
XXVII.5. Tributación estatal: impuestos prediales y al ingreso corporativo 1071
XXVII.6. Otras limitaciones a la regulación estatal del comercio interestatal 1073
XXVII.7. Limitaciones a la jurisdicción personal por el debido proceso 1074
XXVII.8. Transferencias de agua entre cuencas 1075
XXVII.9. La exportación de la pobreza 1076
XXVII.10. Revisiones del conflicto de leyes y de la contaminación 1078
XXVII.11. El poder del comercio 1080
Lecturas recomendadas 1082
Problemas 1083

XXVIII. *La discriminación racial* 1086
XXVIII.1. El gusto por la discriminación 1086
XXVIII.2. La segregación escolar 1088
XXVIII.3. Los requerimientos de la acción estatal 1091
XXVIII.4. Las leyes contra la discriminación 1093
XXVIII.5. La discriminación inversa 1096
Lecturas recomendadas 1100
Problemas 1101

XXIX. *La protección de los mercados libres en las ideas y en la religión* 1103
XXIX.1. La base económica de la libertad de expresión . . . 1103
XXIX.2. El alcance de la actividad protegida: incitación, amenazas, difamación, obscenidad y quema de la bandera 1105
XXIX.3. Restricciones previas, restricciones al punto de vista y el derecho de los foros 1112

XXIX.4. Análisis costo-beneficio de la libertad de expresión: el enfoque por categorías 1115
XXIX.5. La regulación de la radiodifusión 1117
XXIX.6. La publicidad falsa y la relación entre los derechos políticos y los económicos 1119
XXIX.7. El subsidio del discurso 1122
XXIX.8. La economía de la libertad de religión 1123
Lecturas recomendadas 1126
Problemas 1127

XXX. *Cateos, confiscaciones e interrogatorios* 1128
XXX.1. Revisión del derecho a la privacidad 1128
XXX.2. Los remedios de la Cuarta Enmienda 1131
XXX.3. El enigma de la autoincriminación y de confesiones por coerción 1134
Lecturas recomendadas 1137
Problemas 1137

Octava Parte
DERECHO COMPARADO E INTERNACIONAL

XXXI. *Derecho comparado e internacional* 1141
XXXI.1. Derecho comparado 1141
XXXI.2. Los tratados y otros acuerdos internacionales 1144
XXXI.3. El *dumping* y la cuestión del libre comercio 1150
XXXI.4. La inversión extranjera directa 1153
XXXI.5. El Estatuto de Reclamación por Agravios contra Extranjeros 1155
XXXI.6. Las cláusulas de elección de foro o de jurisdicción . . 1156
Lecturas recomendadas 1158
Problemas 1159

Índice analítico 1161
Índice general 1187

La conexión entre el derecho y la economía es profunda y casi inseparable. Pese a esto, muchos especialistas hallaban dificultades para reconocer esa conexión y aplicarla en cuestiones específicas. Por ello, Richard A. Posner presenta un análisis económico de cuestiones legales, correspondientes a un campo relevante, vigente e influyente. Lo que el autor propone es que incluso si el derecho es una doctrina amplia, algunos principios económicos pueden describir gran parte del sistema legal. Por lo tanto, entender la interrelación entre la economía y el derecho es una herramienta para comprender la parte por el todo y el todo por la parte.

Contrario a otros libros de esta índole, Posner abarca en esta nueva edición gran parte del mundo legal y subraya la regulación jurídica de temas fuera del mercado poco comunes como la drogadicción, los robos de obras de arte, los actos sexuales, entre otros, con una lectura más ecléctica y menos técnica y doctrinaria. Además, esta obra se construye alrededor de conceptos legales y no económicos, lo cual sirve para que el derecho se explique como un sistema para el cual la economía resulta de gran utilidad.

FONDO DE CULTURA ECONÓMICA
CONAHCYT
CONSEJO NACIONAL DE HUMANIDADES CIENCIAS Y TECNOLOGÍAS

www.fondodeculturaeconomica.com