

Omar O. Chisari

Compilador

Julio J. Elías

Gustavo Ferro

Daniel Lema

Alfredo M. Navarro

Lucía Quesada

Carlos A. Romero

Christian Ruzzier

PROGRESOS EN MICROECONOMÍA



**ASOCIACION ARGENTINA
DE ECONOMIA POLITICA**

EDICON
FONDO EDITORIAL CONSEJO

 **consejo**

Profesional de Ciencias Económicas de la Ciudad Autónoma de Buenos Aires

Progresos en Microeconomía

Progresos en microeconomía / compilado por Ernesto Rezk. - 1a ed. -
Buenos Aires: Consejo Profesional de Ciencias Económicas de la Ciudad
Autónoma de Buenos Aires, 2011.
242 p.; 21x15 cm.

ISBN 978-987-660-107-8

1. Microeconomía. I. Rezk, Ernesto, comp.
CDD 338.5

ISBN: 978-987-660-107-8

1ra. Edición

Hecho el depósito que marca la Ley 11.723.

Prohibida su reproducción total o parcial por cualquier
medio sin autorización previa del CPCECABA.

EDICON

Fondo Editorial Consejo

Consejo Profesional de Ciencias Económicas de la
Ciudad Autónoma de Buenos Aires

Viamonte 1549 - CABA

Tel. 5382-9200

www.consejo.org.ar

www.edicon.org.ar

Omar O. Chisari

(compilador)

Progresos en Microeconomía

Julio J. Elías
Gustavo Ferro
Daniel Lema
Alfredo M. Navarro
Lucía Quesada
Carlos A. Romero
Christian Ruzzier

EDICON
FONDO EDITORIAL CONSEJO

 **consejo**

Profesional de Ciencias Económicas de la Ciudad Autónoma de Buenos Aires

Sobre los autores

Omar Osvaldo Chisari es Doctor en Ciencias Económicas de la Facultad de Ciencias Económicas, Universidad de Buenos Aires. Es director del Instituto de Economía UADE, universidad de la que fue decano de Ciencias Económicas. Es Investigador Principal del CONICET y Académico Titular de la Academia Nacional de Ciencias Económicas. Fue Presidente de la Asociación Argentina de Economía Política (2004-2006), y entre otras distinciones recibió el Premio Konex (Diploma al Mérito) en 2006, y el Premio Houssay en 1987 y 2003. Entre sus publicaciones, se cuentan varios trabajos sobre equilibrio general computado con aplicaciones al caso de la economía argentina.

Alfredo Martín Navarro es Doctor en Ciencias Económicas egresado de la Universidad de Buenos Aires. Fue Presidente de la Asociación Argentina de Economía Política durante los años 1989 y 1990. Es miembro titular de la Academia Nacional de Ciencias Económicas. Ha sido director de diversos institutos de investigación y ha ejercido la docencia en varias universidades públicas y privadas. Recientemente ha trabajado en temas relacionados con la metodología económica y la neuroeconomía.

Julio Jorge Elías es Doctor en Economía de la Universidad de Chicago y Licenciado en Economía de la Universidad Torcuato Di Tella. Es profesor en el Departamento de Economía y en la Escuela de Negocios de la Universidad del

CEMA (UCEMA). También es director ejecutivo del Centro de Economía de la Creatividad de la UCEMA e investigador asociado del Center of Excellence on Human Capital and Economic Growth and Development de la Universidad de Búfalo. Sus trabajos de investigación recientes se han concentrado en mercados de órganos para trasplantes y en los beneficios no monetarios de la educación. Fue profesor en el departamento de economía de la Universidad de Búfalo durante el período 2005-2008.

Gustavo Ferro es Doctor en Economía por la Universidad del CEMA. Actualmente profesor-investigador del Instituto de Economía de UADE e Investigador Adjunto de CONICET. Sus publicaciones más recientes se relacionan con la economía de mercados regulados (servicios públicos e intermediarios financieros), y su trabajo de docencia se centra en economía de la regulación y finanzas públicas. Ha recibido distinciones nacionales e internacionales por su producción escrita. Ha realizado consultoría para organismos internacionales, entes reguladores y empresas.

Carlos Adrián Romero es Doctor en Economía (UNLP), MSc Economics (Warwick), tiene un Posgrado de Economía (Di Tella), y es Licenciado en Economía (UBA). Es profesor investigador del Instituto de Economía UADE y profesor de Economía Computacional en UADE. Es especialista en economía computacional y microeconomía aplicada. Ha publicado diversos artículos en las áreas de economía de la regulación, defensa de la competencia, CGE y comercio internacional. Ha sido consultor de organismos nacionales e internacionales, fundaciones y empresas privadas.

Daniel Lema es Doctor en Economía en la UCEMA, tiene un Posgrado en Teoría Económica del Instituto Torcuato Di

Tella y es Licenciado en Economía de la Universidad Nacional de Mar del Plata. Es investigador del Instituto de Economía y Sociología del INTA y Profesor en la UCEMA. Ha sido profesor invitado en las Universidades Católica Argentina, Austral, Nacional del Sur, Nacional de Mar del Plata y UBA en las áreas de economía agraria y agronegocios. Ha participado en proyectos de investigación de organismos nacionales e internacionales, tales como FONCYT, CIMMYT, FAO y la European Commission. Ha sido presidente de la Asociación Argentina de Economía Agraria y autor de capítulos de libros y artículos en temas de economía agraria, microeconomía aplicada, organización empresarial y economía política.

Christian Ruzzier es Doctor en Economía (Toulouse School of Economics), Master en Economía Matemática y Econometría (Toulouse School of Economics), Master en Economía (UTDT) y Licenciado en Economía (UADE). Es profesor invitado en la Universidad de San Andrés, donde dicta Economía I en el grado y Economía de las Organizaciones en la Maestría en Economía y en la Maestría en Estudios Organizacionales. También se ha desempeñado como Postdoctoral Fellow en Economía de las Organizaciones en Harvard. Su labor académica ha sido publicada en revistas internacionales, y es asimismo co-editor del libro “Problemas de Defensa de la Competencia en Sectores Regulados en la Argentina”. Ha sido consultor de organismos nacionales e internacionales y empresas de América Latina y Europa.

Lucía Quesada es Doctora en Ciencias Económicas de la Universidad de Toulouse. Es profesora investigadora del departamento de economía de la Universidad Torcuato Di Tella y profesora de microeconomía en UTDT. Es especialista en teoría de los juegos y teoría de contratos. Ha publi-

cado diversos artículos en el área de la microeconomía teórica sobre funcionamiento de mercados de intermediarios de certificación y teoría de contratos aplicada a contratos petroleros.

Índice

Sobre los autores	5
Prólogo	13
Breve Historia de la AAEP	14
Qué es la Asociación Argentina de Economía Política	15
Sede de la AAEP	17
Introducción	19
<i>Omar O. Chisari</i>	
Progresos recientes en Neuroeconomía	25
<i>Alfredo Martín Navarro</i>	
1. Introducción	25
2. El nacimiento y la evolución de la neuroeconomía	25
3. Avances recientes	29
Envejecimiento y decisiones económicas	29
Análisis del altruismo	30
Hábitos y adicciones. Obesidad y alimentación	32
El rol de los neurotransmisores	32
Inversión y <i>animal spirits</i>	33
Rol de las emociones en los procesos decisorios	35
Utilidad cardinal	35
Comparaciones interpersonales de utilidad	36
El modelo de Bernheim	36
Desarrollos en Argentina	38

4. Conclusiones finales	39
5. Referencias bibliográficas	43
Análisis de Mercados No Convencionales: Mercados de Órganos y Repugnancia	49
<i>Julio J. Elías</i>	
1. Introducción	49
2. Introducción de Incentivos Monetarios en el Mercado de Órganos	50
La Situación del Mercado	52
Incentivos Monetarios y el Análisis Económico	54
El Precio de un Riñón para Trasplante	56
Equilibrio de Mercado	59
El rol de la repugnancia en el mercado	61
Bienestar de las personas bajo dos regímenes: Mercado y Prohibición	64
¿Sistema de Mercado?: Principio (Económico) de Compensación versus Voto Mayoritario	66
Cambios en la repugnancia: Efectos de un aumento en la lista de espera renal	68
Lo que es Repugnante depende de las Circunstancias	70
3. Referencias bibliográficas	77
Comparación Teórica y Empírica del Índice de Malmquist y del Indicador de Luenberger. Una aplicación al mercado de seguros	81
<i>Gustavo Ferro - Carlos A. Romero</i>	
1. Introducción	81
2. Tecnología funciones de distancia y medidas de eficiencia	83
Definiciones de tecnología y funciones de distancia	83
3. Descripción de los índices e indicadores de productividad	86
Índice de Productividad de Malmquist	88

Indicador de Productividad de Luenberger	92
Relaciones entre los Índices y los Indicadores	93
4. Aplicación al Sector Seguros de la Argentina	95
Base de Datos	96
Resultados	99
5. Conclusiones	104
6. Referencias bibliográficas	105
Producción y Oferta. Teoría y Aplicaciones en Economía Agropecuaria	109
<i>Daniel Lema</i>	
1. Introducción	109
2. El Enfoque de la Función de Producción	110
3. Un modelo simple de producción	113
4. El Enfoque de Dualidad	118
5. El Análisis de Funciones de Costos	123
6. El Análisis de Funciones de Beneficio	125
7. Funciones de Producción, Heterogeneidad Tecnológica y Productividad	126
8. La Producción en Condiciones de Incertidumbre	130
9. El Enfoque de la Función de Producción Contingente	132
10. Aplicaciones del Enfoque de Producción en Estados Contingentes	141
11. Síntesis y Comentarios Finales	144
12. Referencias bibliográficas	148
Economía de las Organizaciones	153
<i>Christian Ruzzier</i>	
1. Mercado versus organización	153
La pregunta del millón	156
La provisión de incentivos en las organizaciones	159
2. El modelo simple de principal-agente con riesgo moral	161

3. Distorsiones en la provisión de incentivos: la insensatez de pagar por A esperando que se haga B	167
Riesgo y distorsión	175
Múltiples medidas de performance	176
Instrumentos alternativos	179
4. Contratos informales: provisión de incentivos (casi) sin contrato	181
El rol del manager	188
5. Sistemas de incentivos	189
El rol del manager II	192
6. Referencias bibliográficas	193

Intermediarios de certificación y transmisión de información en mercados con información asimétrica 199

Lucía Quesada

1. Introducción	199
2. El modelo	204
Desarrollo temporal del juego	205
3. El equilibrio sin revelación de información	206
Información con valor social: estándares mínimos	211
4. Posibilidad de renegociación e información imperfecta por parte del vendedor	215
Competencia entre intermediarios	222
5. Posibilidad de colusión entre el intermediario y el vendedor	223
El equilibrio sin colusión	227
Equilibrio con un intermediario impaciente	230
Variables de equilibrio como función de la tasa de descuento	233
6. Extensiones y cuestiones pendientes	236
7. Referencias bibliográficas	238

Prólogo

La Asociación Argentina de Economía Política tiene como misión “Promover el análisis económico en el país con miras al adelante de la ciencia”.

En este sentido, sus Reuniones Anuales han sido desde el inicio su principal medio de contribución científica, creando un ámbito estable y adecuado para la presentación de nuestros trabajos, para la evaluación crítica e independiente de la calidad de nuestra producción y para la difusión de nuestros resultados innovadores.

Del mismo modo, las Reuniones Anuales han permitido poner a los asociados, en particular a los jóvenes, en contacto con académicos de primer nivel tanto en el orden nacional como internacional.

A partir de noviembre de 2004 se tomó la iniciativa de reforzar los aspectos de difusión y formación académica incorporando paneles de Progresos en Economía, con la intención de que sean un vehículo eficaz para difundir a toda la comunidad científica los más recientes y destacados avances en cada una de las especialidades de la Economía.

Este libro, que edita ahora la AAEP, con la colaboración editorial del CPCECABA, es el resultado académico del Panel sobre Progresos en Microeconomía desarrollado en la XLV Reunión Anual realizada en la Ciudad Autónoma de Buenos Aires (Universidad Nacional de Buenos Aires) en noviembre de 2010.

Se descuenta que, al igual que lo acontecido con los anteriores paneles, Progresos en Microeconomía será desde ahora una publicación de referencia para nuestros miembros y para docentes, investigadores, graduados y alumnos en Economía.

Ernesto Rezk
Presidente - AAEP

Breve Historia de la AAEP

La AAEP fue fundada en 1957 por los Dres. Juan E. Alemann, Roberto Alemann, Julio Broide, Benjamín Cornejo, Aldo Ferrer, Juan J. Guaresti (h), Carlos C. Helbling, Carlos Moyano Llerena, Julio H. G. Olivera, Federico Pinedo, Oreste Popescu, Ovidio Schiopetto, Francisco Valsecchi y el Ing. Francisco García Olano.

El origen de la AAEP se remonta a sendas invitaciones cursadas por los Dres. Adolph Jöhr y Louis Baudin, a mediados de la década de los cincuenta, a los Dres. Oreste Popescu y Julio H. G. Olivera. Jöhr y Baudin, por entonces pertenecientes a la International Economic Association, sugirieron constituir una asociación representativa de los investigadores en "economía política". La convergencia de iniciativas de los Dres. Popescu y Olivera se cristalizó, el 26 de setiembre de 1957, con la decisión de crear la AAEP. El Dr. Olivera llevó adelante la ejecución de la fundación, la que tuvo lugar formalmente el 18 de noviembre del mismo año.

La historia de la Asociación puede dividirse en tres fases. Durante la primera etapa (1957-1965) la actividad se tradujo en encuentros periódicos para la discusión de temas específicos. En 1958 se realizó la primera reunión de análisis económico. Durante este período la AAEP constituyó varios "centros regionales". La segunda etapa (1966-1972) se caracterizó por la incorporación a la AAEP de representantes de centros e institutos de investigación. A partir de entonces, las reuniones de centros de investigación económica se realizaron en el marco de la AAEP. Se inició en 1968 la rotación de la sede y de las autoridades ejecutivas entre los principales centros. En 1972 tuvo lugar la última reunión de la AAEP organizada sobre la base de trabajos de los centros e institutos de investigación. Desde 1973 hasta el presente la AAEP se encuentra en su tercera etapa, con su sede permanente en la ciudad de Buenos Aires. La AAEP es una institución formada por investigadores y académicos en economía y que interactúa en forma directa con los mismos. El espíritu de una amplia representación institucional y regional ha quedado impreso en la actividad de la AAEP y, en especial, en la práctica de las Reuniones Anuales.

Desde su fundación, la AAEP fue presidida por Julio H. G. Olivera (1957/68), Benjamín Cornejo (1968/70), Víctor J. Elías

(1970/72 y 1978/80), Miguel E. Martínez (1972/74), Horacio Núñez Miñana (1974/76), Aldo A. Arnaudo (1976/78), Rolf R. Mantel (1980/82), Mario L. Szychowski (1982/84), Ana M. Martirena Mantel (1984/86), Luisa Montuschi (1986/88), Alfredo M. Navarro (1988/90), Rinaldo Colomé (1990/92), Juan C. De Pablo (1992/94), Eusebio C. Del Rey (1994/96), Enrique A. Bour (1996/98), José A. Delfino (1998/00), Hildegart Ahumada (2000/02), José Luis Arrufat (2002/04), Omar O. Chisari (2004/06), Alberto Porto (2006/08), Daniel Heymann (2008/10) y Ernesto Rezk (2010/12).

Qué es la Asociación Argentina de Economía Política

La principal actividad de la AAEP es la celebración de una Reunión Anual de discusión de trabajos realizados en el ámbito de la economía, tanto por socios como por no socios, la que tradicionalmente se realiza el mes de noviembre de cada año en distintas ciudades del país, en consulta y cooperación con universidades y centros de investigación económica de la Argentina. En dichas reuniones han participado economistas del país e invitados de otros países, incluyendo prestigiosos economistas extranjeros como Albert Berry, Michael Bruno, Vittorio Corbo, Jacques Drèze, Stanley Fischer, Roger Guesnerie, Arnold Harberger, Hendrik Houthakker, Jean Jacques Laffont, Axel Leijonhufvud, James Mirrlees, F. Modigliani, Marc Nerlove, Luigi Pasinetti, Sherwin Rosen, Pablo Spiller, James Tobin, Wallace Oates, Víctor Volsky, Edward Prescott, T.N. Srinivasan, Finn Kydland, Aloisio Araujo, Stan Metcalfe, Jan Brueckner, Paul Klemperer, Andreu Mas-Colell, Salvador Barberá. Los trabajos son aceptados por una comisión de socios de reconocidos méritos científicos y académicos, designada por el Consejo Directivo.

Los recursos de la AAEP provienen de las cuotas sociales y de los aportes de entidades adherentes. Actualmente la AAEP cuenta con más de 500 socios activos y 17 entidades adherentes: Academia Nacional de Ciencias Económicas (ANCE), Banco Central de la República Argentina (BCRA), Centro de Estudios de Estado y Sociedad (Cedes), Comisión Económica para América Latina y el Caribe (CEPAL), DePablo Consult, Fundación de Investigaciones

Económicas Latinoamericanas (FIEL), Fundación Capital, Instituto Torcuato Di Tella (ITDT), Orlando J. Ferreres y Asociados, Universidad de Buenos Aires (UBA), Universidad del CEMA (UCEMA), Universidad de San Andrés (UdeSA), Universidad Nacional de Córdoba (UNC), Universidad Nacional de Cuyo (UNCu), Universidad Nacional de La Plata (UNLP), Universidad Nacional de Salta (UNSa) y Universidad Nacional del Sur (UNS).

Desde 1964, la AAEP ha llevado adelante 43 Reuniones Anuales en las cuales han sido presentados y discutidos, en un marco de total libertad académica, más de dos mil quinientos trabajos de investigación. Los trabajos presentados en cada Reunión Anual son editados en los Anales de la Asociación Argentina de Economía Política. A estos trabajos hay que sumar los trabajos presentados antes de 1964 y en reuniones científicas, no incluidos en Anales. Los trabajos y comentarios están incluidos también en el web site de la AAEP en Internet. La AAEP no es responsable ni de las opiniones incluidas en los Anales, ni de su protección intelectual.

La AAEP también organiza, en forma no sistemática, reuniones científicas a lo largo del año, en colaboración con otros organismos. Es miembro de la International Economic Association y desde el año 2008 ha lanzado la iniciativa de realizar el ciclo de seminarios de Progresos en Economía con el objeto de contribuir a la creación y fortalecimiento de vínculos con las Universidades que den bienvenida a dichos eventos.

El Consejo Directivo de la AAEP es el órgano de gobierno de la AAEP, y está compuesto de un presidente, dos secretarios y nueve vocales, provenientes de distintos organismos y regiones del país. En 2008 el Consejo Directivo está integrado por: Daniel Heymann (Presidente), Cecilia Rumi (Secretaria), Ramiro Moya (Secretario), Ana María Cerro y Jorge Paz (Secretarios Suplentes), Federico Weinschelbaum, Alfonso Martínez, María Luisa Recalde, Leandro Arozamena, Claudia Nerina Botteon, Andrés López, Marcos Gallacher, Osvaldo Meloni y Fernando Thomé (Vocales). Eran Vocales Suplentes Gabriel Sánchez, Enrique Bour, Daniel Maceira, Jorge Carrera, Graciela María del Carmen García, Laura D'Amato, Eduardo Fracchia, Gustavo Ferro y Adrián Ramos.

Julio H. G. Olivera es Presidente Honorario de la AAEP y los Profesores Albert Berry, Vittorio Corbo, Jacques Drèze, Roger Guesnerie, Arnold Harberger, Jean-Jacques Laffont, Axel Leijonhufvud, James Mirrlees, Marc Nerlove, Alberto Petrecolla, Pablo Spiller, T.N. Srinivasan, Aloisio Araujo, Finn Kydland, Stan Metcalfe, Jan Brueckner, Paul Klemperer, Andreu Mas-Colell y Salvador Barberá han sido declarados Miembros Honorarios de la Asociación.

Sede de la AAEP

Av. Córdoba 637 - 4° piso - (1054) Buenos Aires - Argentina

Tel. (5411) 4314-0246 Fax (5411) 4314-8648

E-mail info@aaep.org.ar

Web-Site de la AAEP en INTERNET:

<http://www.aaep.org.ar>

Introducción

Omar O. Chisari

(Instituto de Economía UADE y CONICET)

Como libros anteriores de Progresos en Economía de la AAEP, la intención de este volumen es informar a los asociados y al público general de los avances recientes en un campo del análisis económico o de la economía aplicada, de modo de bajar los costos de entrada. La idea general de la serie puede resumirse así: “si se tiene que entrar en un campo de la economía de reciente desarrollo, ¿cuáles son sus temas e instrumentos principales?, ¿qué no se debe dejar de leer?”.

Podría decirse que los dos primeros capítulos hacen a los elementos que determinan la construcción y las propiedades de los órdenes de preferencia. Hay aspectos fisiológicos, psicológicos y culturales en juego.

Alfredo Navarro, en el primer capítulo, hace una reseña de qué trata la Neuroeconomía (intersección de economía y neurociencias) y qué avances están disponibles y son de esperar de ella. Como señala Navarro, las observaciones de la *behavioral economics* y algunos casos paradigmáticos de la neurociencia sugirieron que los supuestos macroeconómicos sobre conductas no eran incólumes y merecían una revisión crítica por vía experimental. Era por lo tanto interesante incluir en libro un campo experimental que eventualmente llevará a reexaminar la verosimilitud de los axiomas e hipótesis de comportamiento de los agentes, y ayudará a enten-

der fenómenos todavía bajo discusión, como la forma en que se generan los órdenes de preferencia, los fundamentos del altruismo, la formación de expectativas y la generación de hábitos. Más aún, después de leer el trabajo, surge la inquietud de si la irracionalidad podrá curarse o entenderse algún día.

Los órdenes lexicográficos son un ejemplo de de preferencias a explicar. ¿Se define culturalmente o fisiológicamente qué bienes tienen o no precios subjetivos asociados que pueden transformarlos en mercancías sujetas a transacción? Justamente, sobre en el capítulo que sigue, Julio J. Elías se ocupa de analizar el funcionamiento de los mercados no convencionales, y discute especialmente el rol de la repugnancia para el funcionamiento eficiente del mercado de trasplante de órganos. La definición de repugnancia del artículo se aproxima a la idea de rechazo cultural a realizar transacciones con esos objetos (si bien la prohibición podría surgir de la dificultad de controlar el oportunismo criminal). Y, si ese rechazo es extremo y generalizado entre la población, el mercado simplemente no existirá; aunque esto vaya en contra de la eficiencia económica..

Los dos capítulos que siguen tratan temas de producción.

Gustavo Ferro y Carlos Romero presentan los avances en la medición de productividad a través de números índices e indicadores, un tema de actualidad permanente tanto para los análisis de eficiencia comparada a nivel interno de las firmas, como para la evaluación regulatoria con el objetivo de fijar tarifas. Los discuten en el marco de un ejemplo centrado en el sector asegurador argentino. Analizan las metodologías para estimar cambios de productividad mediante índices de Malmquist e indicadores de Luenberger, y hacen una comparación empírica para ilustrar sus diferencias teóricas. Para ello, utilizan el caso real mencionado del sector de seguros; y la elección del sector no es casual: se trata de

un mercado atomizado, donde la regulación no fija cantidades, y hay también flexibilidad para modificar la mezcla de insumos.

Daniel Lema hace una reseña de la teoría y la estimación de las funciones de producción, con especial énfasis en su aplicación al caso de la producción agropecuaria. Y presenta enfoques recientes de aproximación al caso de la producción bajo una forma especial de incertidumbre, según la cual la cantidad disponible de factores es estocástica. Es un enfoque que merece atención en una variedad de contextos; por ejemplo, puede ayudar a la representación adecuada de los shocks derivados del cambio climático, y contribuir al diseño optimal de instrumentos de adaptación (tanto para la prevención como para la remediación).

El desarrollo de los métodos de benchmarking, que como mencionáramos reseñan Ferro y Romero, respondió en buena medida a la necesidad de tratar empíricamente las asimetrías de información. El reconocimiento de la relevancia de los tipos de asimetrías de información para la toma de decisiones, su definición precisa y las soluciones propuestas para mejorar la eficiencia, han sido uno de los temas dominantes de la microeconomía desde los años setentas. A las aplicaciones y extensiones de la economía de las asimetrías de información se dedican los dos últimos capítulos de este libro.

Christian Ruzzier resume los avances en la economía de las organizaciones. El despliegue de la función de producción, separando de la producción física propiamente dicha de la tecnología organizacional era una tarea pendiente de la microeconomía, en cuanto disciplina básica para explicar buena parte de la administración a partir de la escasez y la distribución de recursos (en este caso, la información). Ruzzier parte de la que él llama “pregunta del millón” implícita en el trabajo de Coase de 1937, referida a por qué se recurre

a las coordinación y los sistemas de incentivos de las empresas si los mercados funcionan bien. Y, admitiendo que esto último no es verdad siempre, avanza hacia la descripción y análisis de modelos de organización interna de las empresas, en tanto instrumentos que no son de primera elección pero de inevitable uso cuando se enfrenta la incapacidad o la ineficiencia de los mercados (*spot*, i.e. de ocasión) para organizar el proceso de producción.

Lucía Quesada discute el rol de los intermediarios de certificación, cuyo principal objetivo (al menos desde el punto de vista normativo) es reducir las asimetrías de información, vía la recopilación de información de los agentes informados y su transmisión al resto de los agentes. Es un trabajo más que oportuno si se tiene en cuenta el papel atribuido a esas instituciones en la crisis financiera actual de la economía mundial. ¿Reducen o aumentan la asimetría esas instituciones cuando son “átomos” de la economía?. ¿Pueden tener comportamientos estratégicos que perjudiquen el óptimo económico desde el punto de vista social? En su trabajo presenta tres aspectos distintos del problema de la transmisión de información por parte de los intermediarios de certificación. Primero, analiza cómo puede beneficiarse el intermediario con la manipulación de la información. Segundo, muestra que la manipulación de la información puede no ser una política óptima si el intermediario es capaz de renegociar el contrato por el servicio de certificación, y que el intermediario puede producir demasiada información, en el sentido de que la disposición a pagar por esa información es menor al costo de obtenerla. Finalmente, investiga los incentivos a la colusión que puede tener el intermediario con el vendedor del producto, con el objetivo de transmitir información falsa que permita vender el producto a un precio más alto.

El método de la microeconomía avanza raudo y se difunde sobre todo el análisis económico. Es difícil defender hoy en día un estudio que no esté firmemente basado sobre el pilar de la consistencia en el comportamiento de los agentes y en el funcionamiento de los mercados.

El progreso de la microeconomía se da también sobre otros frentes, además de esa asimilación de otras disciplinas. Hay avances tanto en el análisis y la economía de sus fundamentos, hacia adentro de la disciplina, como hacia fuera de ella vía las aplicaciones novedosas.

Y es entonces ciertamente difícil resumir en este volumen único todo ese proceso. Las subjetividades de este compilador, las limitaciones del tiempo y del espacio (presupuestarias, al fin) y la amplitud del tema conspiran en contra de una reseña enciclopédica.

De todos modos, la calidad y el interés de los trabajos incluidos en este libro son un paliativo para esas limitaciones. Y cabe esperar que futuros volúmenes de *Progresos en Microeconomía de la AAEP* completen lo faltante.

El compilador y los autores agradecemos a Daniel Heymann y a Ernesto Rezk y a los Consejos Directivos de la AAEP correspondientes la invitación a preparar este trabajo.

Progresos recientes en Neuroeconomía

Alfredo Martín Navarro

(Academia Nacional de Ciencias Económicas)

1. Introducción

El propósito de este trabajo es hacer una reseña y una evaluación de lo que ha ocurrido recientemente en el campo de la neuroeconomía, en especial durante los cinco últimos años. Comenzamos con una breve descripción de las razones por las que aparece esta disciplina y de su evolución durante el corto período que transcurre desde entonces, que apenas llega a menos de dos décadas. En la tercera parte hacemos una descripción de los principales temas que han sido abordados recientemente y por último, en el punto cuarto, realizamos algunas reflexiones sobre el efecto que la neuroeconomía ha tenido y puede tener en el futuro sobre la economía.¹

2. El nacimiento y la evolución de la neuroeconomía

La neuroeconomía es una disciplina que aparece como tal en los primeros años de la década de los noventa, como una conjunción de la economía y la neurociencia. Estas dos ra-

¹ Hemos descripto los avances iniciales en De Schant, Martín y Navarro (2007), Navarro (2007) y Navarro (2008).

mas de la ciencia tuvieron un desarrollo previo de naturaleza tal que hicieron posible que se produjera su confluencia en el estudio de los procesos decisorios.

Los economistas tuvieron en un principio una posición dubitativa sobre si los principios en que se basa la teoría de la elección debían ser adoptados a priori o si debían ser confirmados con experimentos, y si la psicología debía o no ser tenida en cuenta al formularse los principios básicos de la teoría microeconómica. Edgeworth creía que en algún momento la tecnología permitiría que contáramos con un aparato que llamaba *hedonímetro* que serviría para medir la utilidad, y Keynes incorpora conceptos originados en la psicología, como el de *animal spirits*. Posteriormente, a partir del final de la década de los años treinta del siglo pasado, siguiendo las ideas, entre otros, de los economistas de la escuela austríaca, y también de Robbins, Pareto y Samuelson entre otros, se descartó totalmente la comprobación empírica de los principios económicos básicos, se supuso que los agentes económicos actuaban racionalmente² y se dejó de lado la psicología como elemento útil para el análisis de los procesos decisorios.³ Esta actitud comenzó a ser cuestionada por los economistas que integraban la corriente conocida como *behavioral economics*, quienes con base principalmente en los trabajos de los años ochenta de Kahneman y Tversky, comenzaron a reclamar una mayor presencia de la psicología,

² En De Schant, Martín y Navarro (2007) hemos realizado un análisis detallado de la evolución del concepto de racionalidad en la teoría económica en relación con esta cuestión. También puede encontrarse una descripción del funcionamiento de la técnica conocida como fMRI.

³ Sin embargo no había unanimidad y desde un principio se alzaron voces críticas. Por ejemplo, Hutchison (1938) cuestiona el método que propone Robbins y considera que todos los principios básicos en los que se asienta la economía, deben pasar por el proceso de falsificación de K. Popper, porque de lo contrario, al aplicarse el criterio de demarcación de este autor, la economía no podría ser considerada como una ciencia.

y lograron que la experimentación se propagara con fuerza en el campo de la economía.⁴

Por otra parte, los estudiosos de la neurociencia comenzaron a preocuparse también por los procesos decisorios y por los efectos de las alteraciones en el cerebro sobre la conducta humana. En un principio sólo pudieron analizarse algunos casos especiales, como el de Phineas Gage, quien en 1848 sufrió un accidente que le destruyó una parte del cerebro, lo que permitió estudiar los cambios en su conducta originados en ese hecho. Luego estudiaron la conducta de los animales, especialmente de los simios, hasta que más recientemente aparecieron diversas técnicas, que por no producir daños ni lesiones, permitieron visualizar algunos cambios que suceden en el interior del cerebro en el momento de la toma de decisiones, y después de probarse con varios métodos de análisis, se desarrolló y difundió la técnica que se conoce como *resonancia magnética funcional* (fMRI), que consiste en un procedimiento totalmente inofensivo por el cual se pueden obtener imágenes del cerebro en el momento en que se toman las decisiones, las que permiten determinar cuáles son las áreas que se activan en ciertas circunstancias, debido a la mayor afluencia de sangre.⁵ Por otra parte, los estudios sobre los neurotransmisores permiten establecer su influencia sobre la conducta humana y sobre las decisiones, como ocurre, por ejemplo con la dopamina, la serotonina y la oxitoxina. Los neuro-

⁴ Una excelente descripción de la teoría del comportamiento económico, de sus orígenes, sus métodos y el estado actual de esa corriente puede verse en Camerer, Loewenstein y Rabin (2004).

⁵ Recientemente se han comenzado a utilizar las técnicas conocidas como Transcranial Magnetic Stimulation (TMS) y Direct Current Stimulation (tDCS), que consisten en utilizar electrodos que al activar o desactivar determinadas áreas del cerebro permiten analizar el comportamiento de los agentes económicos ante ese tipo de cambios en la actividad cerebral. Una aplicación de estas técnicas puede verse en Fehr (2009).

científicos tomaron conocimiento de la existencia de la teoría económica y descubrieron su riqueza, con lo que ambas disciplinas confluyen en una actividad en la participan neurocientíficos y economistas.

Estos métodos comenzaron a utilizarse hacia 1992, y a partir de entonces se realizaron las primeras reuniones científicas, mientras que las revistas especializadas, tanto las relacionadas con la neurociencia como las económicas, publican los primeros trabajos sobre estos temas. En el año 2003 se creó la *Society for Neuroeconomics*, cuyo primer presidente fue Paul Glimcher, que desde entonces organiza reuniones anuales. En 2005 se creó la *Association for NeuroPsychoEconomics*, que edita la revista *Journal of Neuroscience, Psychology and Economics*, que realiza también reuniones periódicas, y en noviembre del 2009, la *Asociación Española de Neuroeconomía*. Durante el último quinquenio el crecimiento ha sido explosivo, comenzaron a aparecer publicados centenares de trabajos científicos y se crearon varios centros de investigación.⁶ El trabajo publicado en el *Journal of Economic Literature*, en 2005 de Camerer *et al.*, un *survey* que cuenta lo avanzado hasta ese momento, despertó un marcado interés entre los economistas, entre los que me cuento. En el año 2009 apareció la obra de Glimcher *et al.*, *Neuroeconomics: Decision Making and the Brain*, que recoge la mayoría de los estudios más relevantes publicados en torno a esta cuestión y pone en las manos de todos los interesados trabajos publicados en revistas que, por pertenecer a otra disciplina, no son conocidas por todos

⁶ *ScienceDirect* en 20 revistas registra 56 trabajos sobre neuroeconomía en 2006 y ese número se quintuplica en 2009, lo que da una idea del crecimiento exponencial del esfuerzo de los economistas y especialistas en neurociencia que se dedican a esta disciplina y el interés con que este tipo de trabajos es recibido.

los que desde la economía o la neurociencia se interesan en estos temas.

Recientemente, dos revistas importantes dedicaron un número completo a la neuroeconomía: en 2008, *Economics & Philosophy*,⁷ y en 2010, *Journal of Economic Methodology*,⁸ en los que se discuten las implicancias metodológicas de la neuroeconomía, y la pregunta básica es ¿debe la economía abrir la supuesta caja negra que es el funcionamiento del cerebro, o más bien dejar este tema para la neurociencia y acotar su campo a los problemas decisorios, sin entrar a analizar cuestiones más profundas, como el funcionamiento del cerebro, la interacción entre sus distintos componentes o el efecto de los neurotransmisores sobre la conducta humana?

3. Avances recientes

Durante los últimos años aparecieron una gran cantidad de trabajos en los que se aplican avances en neurociencia a problemas económicos. Si bien es imposible hacer un detalle exhaustivo, vamos a referirnos a los que consideramos de mayor interés para el avance de la microeconomía.^{9, 10}

Envejecimiento y decisiones económicas

La teoría económica, cuando estudia los procesos decisorios de los agentes económicos, no tiene en cuenta la edad de quien decide. Un trabajo de Mohr, Li y Heekeren (2010) analiza los cambios en las decisiones de las personas en fun-

⁷ N° 24, 2008.

⁸ Vol. 17, N° 2, 2010.

⁹ Se han realizado también numerosos estudios en los que se analizan problemas de marketing y de gerenciamiento, pero consideramos que caen fuera del campo de lo que conocemos como neuroeconomía -que se limita al análisis de problemas económicos-, y pertenecen a lo que Hubert (2009) denomina *neuromarketing* o *neurociencia del consumo*.

¹⁰ El orden en que son tratados los diferentes temas es aleatorio.

ción de su edad, y llega a la conclusión de que neurotransmisores como la dopamina y la serotonina son producidos en menor cantidad a medida que vamos envejeciendo, y eso tiene efectos sobre la apreciación de las recompensas, del riesgo asumido y de la espera por la recepción de la recompensa. Los mencionados autores realizan experimentos donde se puede observar, mediante el uso de fMRI, que ciertas áreas del cerebro, como el striatum ventral y partes de la corteza prefrontal, se activan de distinta forma en los jóvenes y en los adultos. Asimismo analizan los cambios que se producen en las respuestas cuando se administran drogas que activan la producción de esos neurotransmisores, y llegan a la conclusión de que los adultos son menos propensos a aceptar riesgos. Por otra parte, la tasa de descuento en las recompensas diferidas es superior en las personas de edad intermedia, pero mayor en los adultos que en los jóvenes.

Esta línea de investigación se vuelve importante en la medida en que las sociedades van envejeciendo como consecuencia del aumento de la esperanza media de vida al nacer y, si bien recién está en sus comienzos, es un campo con perspectivas de gran interés.

Análisis del altruismo

Adam Smith suponía que los seres humanos actuaban movidos por el propio interés y por la simpatía, pero posteriormente la economía fundamentó el análisis económico en el supuesto de que los agentes económicos actúan en forma egoísta, y este principio permitió fundamentar razonablemente la microeconomía. Sin embargo, luego se comenzó a estudiar el comportamiento altruista, que se encuentra entre algunos animales, como en ciertos insectos y en los animales más desarrollados, pero que alcanza características superiores entre los seres humanos. Fehr y Fischbacher (2003) sostienen que las relaciones sociales y la organización de la

sociedad están ligadas al concepto de altruismo, que está muy heterogéneamente distribuido entre los seres humanos: algunos son predominantemente egoístas, mientras que otros tienen una mayor propensión a actuar desinteresadamente. Concluyen que las actitudes egoístas generan respuestas de la misma naturaleza, y que las actitudes altruistas tienen el mismo efecto, pero también reconocen la importancia del factor genético en el comportamiento.

Si bien el comportamiento totalmente egoísta es una buena simplificación en el funcionamiento de los mercados, en las interacciones entre los individuos pareciera ser que la función de utilidad por maximizar es de la siguiente forma:

$$(1) \quad U_i = x_i + \sum_j v_j (\kappa_j^i) \cdot x_j,$$

donde U es la utilidad del individuo i , x la utilidad egoísta del individuo i , v es la utilidad que significa para el individuo i , la utilidad de los otros j individuos que componen el grupo social, y κ , el grado de simpatía que el individuo i tiene hacia el individuo j , lo que puede tener signo positivo o también negativo (en el caso de que j tenga un ingreso demasiado elevado o se le quiera sancionar, justa o injustamente por alguna de sus acciones pasadas).

Como los individuos se enfrentan a un *trade off* entre la utilidad que proviene de su propio consumo de bienes o servicios y el de las otras personas, debemos pensar en una situación de equilibrio, que varía mucho entre los distintos individuos. Dado que la solidaridad es un elemento relevante en la convivencia social, el estudio de sus fundamentos básicos es de sumo interés para las ciencias sociales. En el trabajo referido los autores encuentran un paralelismo entre los centros cerebrales que se activan en respuesta a la satisfacción egoísta, y los que reaccionan ante la que produce el bienestar de las demás personas. Los cambios en la conducta debidos al suministro de oxitocina o la activación o desactiva-

ción de determinadas áreas mediante el TMS permiten avanzar hacia las bases neuronales del comportamiento social.

Hábitos y adicciones. Obesidad y alimentación

Los hábitos y las adicciones condicionan la elección de los agentes económicos, y la neuroeconomía ha realizado aportes que permiten conocer cómo funcionan los mecanismos del cerebro para que condicionen la conducta y por tanto la elección de los bienes consumidos. Esto permite analizar el consumo de sustancias tóxicas, como las drogas y el alcohol y determinados hábitos alimenticios que conducen a la obesidad y a perturbaciones en el organismo debidas a la ingesta de una alimentación inadecuada. Avena *et al.* (2009) analizan los mecanismos que conducen a la ingesta de alimentos ricos en grasas y azúcar, mientras que Bickel *et al.* (2007) analizan la drogadicción mediante el estudio de la forma en que se activan determinadas partes del cerebro en el proceso del análisis de la preferencia intertemporal.

El rol de los neurotransmisores

Paul Zak ha trabajado desde hace varios años en la relación entre determinados neurotransmisores y la conducta humana y recientemente presenta los avances logrados en esta línea de investigación en Vercoe y Zak (2010), donde describen un experimento en el que el sujeto A dispone de una cantidad de recursos, de los cuales puede transferir una parte al sujeto B, quien atraviesa por una situación difícil o desagradable. Este último puede rechazar la transferencia si la considera incorrecta; en ese caso, A es sancionado perdiendo los recursos que le habían sido otorgados.

En este experimento los autores analizan el efecto de la oxitocina, cuyo nivel controlan haciendo aspirar una sustancia que la contiene a las personas que participan del experimento, y afirman que la oxitocina ayuda a la secreción de

dopamina y serotonina, de las cuales la primera hace que la ayuda a los demás se convierta en algo placentero, mientras que la segunda mejora el estado de ánimo y reduce la ansiedad que produce la presencia de extraños. Finalmente experimentan elevando los niveles de oxitocina, y comprueba que esto produce un comportamiento más proclive a la ayuda a las demás personas, mientras que una elevación de los niveles de testosterona tiene el efecto contrario.

Inversión y animal spirits

Uno de los campos que aparece como más fértil para la neuroeconomía es la teoría de la inversión, que continúa en la actualidad los esfuerzos incipientes de Keynes en los años treinta, cuando advirtió que era necesario analizar la inversión desde la perspectiva de la psicología.

Schwartz (2010) realiza un análisis del libro de Akerman y Shiller (2009) sobre la última crisis, en el que sostienen que para interpretar lo sucedido debe tenerse presente el concepto keynesiano de *animal spirits*.¹¹ Se pregunta por qué los autores no han recurrido a los aportes de la economía del comportamiento y de la neuroeconomía, y en cambio sostienen, a mi juicio acertadamente, que este concepto debe analizarse desde la perspectiva de la microeconomía. La aplicación del principio de racionalidad deja de lado este elemento de la conducta de los agentes económicos, que requiere de un tratamiento distinto: Keynes se limita a destacar su presencia en la conducta humana, que tiene la acción como un objetivo, un deseo de hacer cosas y emprender tareas que si fueran evaluadas racionalmente, tal vez no se llevarían a cabo, y esos deseos de actuar no son constantes, sino que tienen más fuerza en algunos períodos y menos en otros. Si

¹¹ Koppl, (1991) realiza un detallado análisis de este concepto, que ayuda a valorar la importancia del aporte original de Keynes.

utilizáramos los avances de la neurociencia relacionados con el rol de las emociones y la aversión al riesgo y a las pérdidas, podríamos ir un poco más lejos y determinar los mecanismos que subyacen en la conducta y nos impulsan a actuar en algunas circunstancias y no en otras.

Los agentes económicos presentan aversión o adicción al riesgo, y de esta forma adoptan conductas que no son óptimas. El trabajo de Kuhnen y Knutson (2005) analiza mediante la técnica de fMRI la forma en que se activan distintas partes del cerebro en el proceso de toma de decisiones, y realizan un experimento, cuyos resultados luego analizan estimando un modelo econométrico tipo *logit*, arribando a la conclusión de que la activación del *núcleo accumbens* está asociada con comportamientos riesgosos, y la *ínsula anterior* con conductas que muestran renuencia al riesgo. Esta línea de investigación parece abrir un camino promisorio, ya que por esa vía tal vez podamos aproximarnos al concepto de *animal spirits*, y comprender mejor los mecanismos internos que se asocian con la inversión y con el funcionamiento de los mercados de valores o bienes.

Un trabajo reciente, de Bermejo, P., Dorado, R., Zea-Sevilla, M. y Sánchez Menéndez, V. (2010) sostiene que en el cerebro humano existen dos sistemas complementarios: el sistema de recompensa cerebral y el de aversión a la pérdida. Lo novedoso de esto es que ambos sistemas tienen un pie en la *corteza* y otro en la *parte límbica* del cerebro. El primero está integrado por el *núcleo accumbens*, que se activa cuanto mayores son los beneficios económicos, y la *corteza prefrontal*, que se activa cuanto mayor es la *probabilidad* de tales beneficios. El segundo está menos definido pero pareciera comprometer la *amígdala*, la *ínsula* y el *hipotálamo*. En situaciones normales los agentes económicos son en alguna medida renuentes al riesgo, pero lesiones en el cerebro pueden afectar sus procesos decisorios, haciéndoles

más proclives a la toma de riesgo. De esta manera, el análisis de alteraciones en distintas partes del cerebro, podría ayuda a comprender el rol de cada una de sus partes. Citan un experimento realizado con monos capuchinos en la Universidad de Yale, donde pudieron comprobar que en su proceso decisorio ponen de manifiesto aversión por el riesgo. La decisión se adopta luego de la interacción de ambos sistemas, dependiendo de cada individuo y de la forma como es afectada cada uno de ellos. Luego analizan distintas patologías, tales como la depresión y la manía, el alcoholismo, la ludopatía y la enfermedad de Alzheimer y llegan a la conclusión de que la forma en que toman sus decisiones las personas afectadas por estas enfermedades difiere de las personas normales, lo que en sirve de ayuda para comprender los procesos de estas últimas.

Rol de las emociones en los procesos decisorios

Los procesos decisorios son vistos desde una perspectiva diferente, continuando con la línea de pensamiento de Ainslie (1991), quien sostiene que dentro del cerebro humano existen diversos centros de decisión que compiten entre sí. Mientras una parte de mi cerebro me sugiere consumir alcohol, otra me induce a abstenerme, y es de la interacción de ambas fuerzas de donde saldrá una resultante que definirá el camino a seguir. Kincaid y Ross (2009) van un poco más lejos y sostienen que la racionalidad no está en los agentes económicos, sino que se encuentra alojada en grupos de neuronas, que no necesariamente coinciden con los demás componentes del cerebro.

Utilidad cardinal

Los economistas utilizamos el concepto ordinal de la utilidad, que hace que no vayamos más allá de la comparación de las utilidades de dos bienes, para suponer que el que pre-

ferimos es el que nos proporciona mayor utilidad, ya que esta no se puede medir. Sin embargo, Stuphorn (2006) describe experimentos con simios, en los cuales intenta demostrar que la utilidad es medible. Si esto se comprobara en los seres humanos tendría implicaciones muy relevantes en la microeconomía.

Comparaciones interpersonales de utilidad

Kahneman (2009) plantea un problema que ha preocupado durante años a los economistas: ¿son posibles las comparaciones interpersonales de utilidad? Su respuesta es que si los cerebros son suficientemente parecidos en sus estructuras y en sus funciones como indican las correlaciones encontradas entre la actividad cerebral y determinados estados psicológicos, no pareciera justificar demasiadas diferencias entre ellos, y eso cuestionaría la justificación de que las comparaciones interpersonales de utilidad hayan sido proscriptas de manera absoluta.

El modelo de Bernheim

Bernheim (2009) realiza un análisis profundo y atractivo sobre la utilidad de la neuroeconomía para el enriquecimiento de la microeconomía desde la perspectiva de un modelo de raíz econométrica. Propone para el análisis del efecto de las variables neuroeconómicas, un modelo del siguiente tipo:

$$(2) \quad \mathbf{y} = Y [Z(\mathbf{x}, w), \mathbf{x}, w]$$

donde \mathbf{y} es un vector de bienes que compra un individuo, el cual depende de variables que podríamos llamar “convencionales”, \mathbf{x} , tales como los impuestos, el precio, el ingreso, el género y la edad, entre otras, y variables que tienen el carácter de neurológicas, que están representadas por Z , que es función de $Z(\mathbf{x}, w)$, por lo que están influidas por las propias variables exógenas \mathbf{x} , y por w , que constituye un

término de ruido. La neuroeconomía tiene sentido si el pronóstico o la explicación de y mejoran con la explicación de estas variables. Bernheim plantea una cuestión muy interesante cuando se pregunta hasta donde debemos llegar en nuestro análisis retrospectivo. Sugiere que si seguimos buscando causas de la conducta en cuestiones más alejadas estaremos dentro de un proceso similar al de las muñecas rusas y los economistas terminaremos estudiando teoría de las cuerdas. A veces, dice, depende de cuál sea nuestro propósito, no es necesario llegar a los fundamentos. Newton explicó por qué la manzana cae hacia la tierra, y con su teoría podemos calcular con gran precisión el tiempo que mediará entre el momento en que se desprende del árbol y aquel en el que toca el suelo, pero no nos dice por qué los cuerpos se atraen unos a otros. En primer término Bernheim analiza el rol de la neuroeconomía en la economía positiva, para lo que divide las variables en exógenas y endógenas. En esta cuestión es escéptico, y si bien reconoce que si pudiéramos realizar un estudio del estado cerebral de una muestra de personas antes del funcionamiento del mercado, podríamos tener una idea de si el comportamiento de ese mercado va a ser agresivo o cauteloso, en general su trabajo está lleno de objeciones al uso de la neurociencia como fundamento de la economía y finaliza diciendo:

I do not mean to trivialize neuroeconomic studies that establish correlations between endogenous brain activity and choices. Findings that help us understand the neurobiology of cognition and decision making have unquestioned scientific merit. I am concerned, here, only with a narrow issue, whether those findings illuminate traditional economic questions. For all its scientific merit, the ability to predict choices from endogenous brain activity is largely orthogonal to the current objectives of the mainstream economics.

Queda pendiente, dice, que se siga desarrollando esta disciplina para poder hacerse un juicio definitivo, pero no cree que vaya a tener un efecto revolucionario sobre la teoría económica contemporánea¹².

Otros temas que consideramos de interés y que la literatura ha analizado durante los últimos años señalando progresos que creemos importantes son las cuestiones relacionadas con la aversión al riesgo (Rick, 2010) y la preferencia intertemporal (Benhabid et al., 2010 y MacKeil Carter, 2010), entre otros.

Desarrollos en Argentina

En nuestro país recién comienzan los estudios sobre neuroeconomía, y si bien existe un desarrollo destacado en el campo de la neurociencia, hasta donde tengo conocimientos se ha avanzado relativamente poco. Pero dentro de los trabajos realizados son de destacar los aportes realizados por Federico Contigiani (2010), quien en su trabajo de tesis incorpora tres estudios sobre neuroeconomía, que realiza desde una perspectiva matemática. Por una parte, estudia el problema del consumo de bienes adictivos, siguiendo las ideas expresadas en Gavrila *et al.* (2005), donde demuestra que es posible que ante una restricción presupuestaria el adicto tenga varios equilibrios posibles, o que está influido por la historia de su comportamiento anterior y de qué manera el individuo reacciona ante distintos impulsos del ambiente. En el segundo trabajo formaliza la relación entre los sistemas emotivo o automático y cognitivo o lento, y en el tercero analiza mediante la técnica de redes neuronales los comportamientos ante preferencias intertemporales, gene-

¹² La revista *American Economic Review Microeconomics* del mes de agosto de 2009, además de incorporar este trabajo, publica otros tres que se refieren al modelo de Bernheim, realizados por Gul y Pesendorfer (2009), Rustichini (2009) y Sobel (2009), que realizan interesantes observaciones que no desarrollamos por falta de espacio.

rando un modelo matemático que da fundamento a los hallazgos en esta área de Loewenstein y Prelec (1992). Por otra parte, Caramuta, Contigiani y Thomé (2010) presentan un modelo donde se presume que el agente económico utiliza su memoria para realizar sus elecciones, evaluando sus decisiones sobre la base de la suma de las utilidades que obtuvo en el pasado.

4. Conclusiones finales

Según Fumagalli (2010), existen tres actitudes ante la neuroeconomía entre los economistas. La primera es la de los *entusiastas*, que estarían representados por Camerer *et al.* (2005), y, entre otros, Barkley Rosser *et al.* (2010), quien nos muestra su opinión favorable:

For me, at least, neuroeconomics, even if about 5% is directly relevant for economists, has an important message – and it doesn't seem to be the message that people are conveying. People has the idea that you are going to be able to relate specific parts of the brain to different specific decisions, a sort of twenty-first century phrenology....For me what is important is the notion that the brain itself is a complex system, and, in some moods, other networks get activated, and then in other moods, other networks get activated and we may switch from one to another....Neuroeconomics really tell us that people behave in different ways, and possibly take contradictory decisions in different circumstances.

La segunda es la de los *escépticos*, cuya manifestación más relevante es el trabajo de Gul y Pesendorfer (2008), quienes entienden que la neuroeconomía está fuera del campo de estudio de la economía, que las imágenes sobre la activación de determinadas partes del cerebro no añaden nada al conocimiento de los economistas y que poco o nada puede esperarse de esta disciplina. “La evidencia producida por la neurociencia no puede refutar los modelos económi-

cos porque estos no estudian la fisiología del cerebro”, afirman en forma categórica y terminante.

Por último menciona a los *moderados*, que sin compartir el optimismo de los primeros, aceptan que la neuroeconomía puede ofrecer aportes interesantes a la fundamentación de la microeconomía, pero con ciertas limitaciones. Un ejemplo de esta posición es la de Harrison y Ross (2010), quienes diferencian entre lo que denominan *behavioral economics in the scanner (BES)* y la *neurocelular economics (NE)*. Esta última es la línea de investigación que propone Glimcher (2003), que estaría fuera del alcance de las críticas de Gul y Pesendorfer, ya que elabora modelos matemáticos que explican el comportamiento y que luego son sometidos al análisis empírico con los métodos de la neuroeconomía.¹³ En este trabajo se realiza una evaluación de las contribuciones que la neuroeconomía ha producido sobre la teoría económica y de las perspectivas futuras de esta disciplina, de manera formal e imparcial. La neuroeconomía incorpora al comportamiento de los agentes económicos variables que hasta ahora se consideraban simplemente ruido, y suponiendo que estaban ocultas al análisis, consideran que estamos abriendo cajas negras. Discuten la conveniencia de abrir esas cajas negras o muñecas rusas. Es preciso entonces saber qué camino debemos seguir. También consideran que en la economía normativa las posibilidades son mayores que en la economía positiva, pero en todo momento reconocen la conveniencia de continuar con esta línea de estudio, aunque con la cautela suficiente, al menos hasta ver con más claridad los frutos de esta línea de trabajo.

En esta misma línea, Harrison (2008) afirma que considera que la neuroeconomía puede brindar aportes interesantes en el futuro, pero cambiando la forma en que se ha venido

¹³ Ver un análisis de esta línea de investigación en Navarro (2007) y una descripción actualizada de esta línea de pensamiento en Ross (2008).

desarrollando en sus principios. Critica la forma en que se realizan los experimentos¹⁴ y que lo avanzado hasta ahora aporta poco o nada al conocimiento que los economistas ya tenían sobre los procesos decisorios.

¿Es la neuroeconomía una rama de la teoría del comportamiento? La mayoría de los autores consideran que la respuesta es afirmativa, ya que ambas utilizan experimentos para comprobar las hipótesis, lo que es diferente a la forma de trabajar de la microeconomía neoclásica, que partía de una serie de supuestos básicos que no podían ser objeto de comprobación empírica. Sin embargo debemos tener en cuenta, en mi opinión, que la psicología y la neurociencia analizan objetos que se encuentran en dos mundos distintos de acuerdo con la metafísica de Popper: mientras la teoría del comportamiento estudia objetos de mundo 2, la neuroeconomía va a un estrato más profundo, a lo físico, a las células y las neuronas, que forman parte del mundo 1.

Supongamos que aceptamos que la neuroeconomía solamente nos dice qué zonas del cerebro están más activas en los procesos decisorios y eso poco o nada nos aporta para que tengamos una mejor teoría económica. Aun en ese caso valdría la pena estudiar neurociencias porque eso aumentaría nuestros conocimientos sobre los seres humanos, cuya conducta es el objeto de estudio de la economía y además nos proporcionaría el placer intelectual del conocimiento científico, de la misma manera que si estudiáramos física cuántica o los fundamentos genéticos de la biología. En ese caso, estaríamos incorporando la neurociencia a la lista de

¹⁴ Critican los reducidos tamaños de las muestras utilizadas, lo que está originado en lo costoso de algunos experimentos. De todas maneras esta crítica solamente afecta a la mayoría de los trabajos que se han publicado, pero puede fácilmente solucionarse en el futuro mejorando los procedimientos estadísticos y econométricos utilizados para evaluar los resultados de los experimentos.

conocimientos que además de los económicos debería tener un economista, que enumera Keynes (1924) en su conocido obituario de Marshall, en el que dice que los economistas deberían tener un cierto dominio sobre cuestiones relacionadas con matemática, historia, ciencia política y filosofía entre otras disciplinas. ¿Sería posible imaginar un economista que careciera de esos conocimientos, a pesar de que no caen estrictamente dentro del campo de la teoría económica?

En el presente trabajo hemos detallado algunas de las hipótesis que han surgido de esta nueva disciplina y que podrían tener como consecuencia que dispusiéramos de una valiosa herramienta para analizar la conducta humana desde una perspectiva más profunda. Cuanto más conozcamos la naturaleza humana estaremos en mejores condiciones de elaborar políticas que mejoren el bienestar de la sociedad, y tal vez la neuroeconomía pueda ayudarnos este sentido.

Pero para que nos sirva para entender mejor el funcionamiento del sistema económico o para diseñar la política económica de manera que incremente el bienestar de los seres humanos queda por delante un largo camino que, sin embargo, ha comenzado a recorrerse, pero solamente el transcurso del tiempo nos permitirá develar la incógnita respecto del valor de sus aportes y de la persistencia de los economistas en su interés por este tipo de estudios.¹⁵

Es también necesario analizar el rol de los economistas en la neuroeconomía. Es una disciplina donde los estudios del funcionamiento del cerebro deben ser realizados por especialistas en neurociencia y los economistas tienen la misión de plantear los problemas a ser estudiados, y en base a los avances que se realicen diseñar modelos económicos en base a los nuevos conocimientos que se adquieran sobre

¹⁵ Esto es coincidente con lo que sostienen Craver y Alexandrova (2008) en el sentido de que la neuroeconomía recién comienza su tarea y existen muchas ideas imprecisas que van a ir tomando forma paulatinamente.

los procesos decisorios. A medida que la neurociencia nos vaya mostrando la manera en que funciona el cerebro en los procesos decisorios, los economistas estaremos en condiciones de diseñar modelos que se aparten del criterio estricto de la racionalidad¹⁶ y avanzar en el conocimiento del comportamiento de los agentes económicos y de los distintos tipos de mercados.

5. Referencias bibliográficas

- Ainslie, G. (1982) *Picoeconomics*. Cambridge: Cambridge University Press.
- Akerlof, G. y Shiller, R. (2009) *Animal Spirits*. Princeton: Princeton University Press.
- Arrow, K. (1987) "Economic theory and the hypothesis of rationality", en *The New Palgrave*. Londres: The Macmillan Press Limited.
- Avena, N., Rada, P., Hoebel, B. (2009) "Sugar and fat bingeing have notable differences in addictive-like behavior". *The Journal of Nutrition* 139(3), pp. 623-628.
- Barkley Rosser, J., Holt, R. y Colander, D. (2010) *European Economics at a Crossroad*. Cheltenham: Edward Elgar Pub.
- Barnea, A., Cronkvist, H. y Siegel, S. (2010) "Nature or nurture: What determines investment behavior?" *Journal of Financial Economics*. Agosto.
- Benhabib, Jess & Bisin, Alberto & Schotter, Andrew, (2010) "Present-bias, quasi-hyperbolic discounting, and fixed costs," *Games and Economic Behavior*, Elsevier, vol. 69(2), pages 205-223, July.

¹⁶ Arrow (1987) sostiene que es posible diseñar modelos económicos que donde se suponga comportamientos no racionales.

- Bermejo, P., Dorado, R., Zea Sevilla, M. y Sánchez Menéndez, V. (2010) “Neuroanatomía de las decisiones financieras”. *Neurología*. doi: 10.1016 / j.nrl.2010.09.015.
- Bernheim, C. D. (2009) “On the potential of Neuroeconomics: a Critic (but hopeful) Appraisal”. *American Economic Journal Microeconomics*. Vol 1. N° 2.
- Bickel, W., Miller, M., Yi, R., Kowal, W., Lindquist, D. y Pitcock, J. (2007) “Behavioral and neuroeconomics of drug addiction”. *Drug and Alcohol Dependence*. Vol 90. Suplemento 1. pp. 585-591.
- Camerer, C., Loewenstein, G. y Rabin, M. (2004) *Advances in Behavioral Economics*. Princeton: Princeton University Press.
- Contigiani, F. (2010) *El efecto de los estímulos ambientales*. Tesis Doctoral. Universidad Nacional del Sur.
- Contigiani, F. (2010) *La dinámica de los sistemas afectivo/cognitivo en el proceso de toma de decisiones: el enfoque de Metcalfe y Mischel desde la óptica de la neuroeconomía computacional*. Tesis Doctoral. Universidad Nacional del Sur.
- Contigiani, F. (2010) *Un modelo neurocomputacional de descuento hiperbólico en problemas de elección intertemporal*. Tesis Doctoral. Universidad Nacional del Sur.
- Craver, C, y Alexandrova, A. (2008) “No revolution necessary”. *Economic & Philosophy*. Vol. 24, N° 3.
- De Schant, F., Martín, J. y Navarro, A. (2007) “Neuroeconomía y metodología: algunas reflexiones iniciales”. *Anales de la Academia de Ciencias de Buenos Aires*, www.a-navarro.com.ar.
- Fehr, E. y Fischbacher, U. (2003) “The nature of human altruism”. *Nature*. Vol 425, pp 785-793.

- Fehr, E. (2009) "Social Preferences and the Brain". En Glimcher, P., Camerer, C., Fehr, E. y Poldrack, R (eds.). *Neuroeconomics: Decision Making and the Brain*. Londres: Elsevier
- Fumagalli, H. (2010) "The disunity of neuroeconomics: a methodological appraisal". *Journal of Economic Methodology*. Vol 17. N° 2.
- Gavrila, C., Feichtinger, G., Tragler, G., Hartl, R. y Kort, P. (2005) "History-dependence in a rational addiction model", *Mathematical Social Science*. 49, pp. 273-293.
- Glimcher, P. (2003) *Decisions, Uncertainty, and the Brain*. Cambridge: The MIT Press.
- Glimcher, P. W. (2009) "Neuroeconomics: History". *Encyclopedia of Neuroscience*, pp. 285-290.
- Glimcher, P. W., Camerer, C., Fehr, E. y Podrack, R. (2009) "A Brief History of Neuroeconomics". En *Neuroeconomics: Decision Making and the Brain*. Londres: Elsevier.
- Gul, F. y Pesendorfer, W. (2008) "The case for mindless economics. En *The handbook of economic methodology*, A. Caplin y A. Schotter (eds.) Oxford: Oxford University Press.
- Gul, F. y Pesendorfer, W. (2009) "A Comment on Bernheim Appraisal of Neuroeconomics" *American Economic Journal Microeconomics*. Vol 1. N° 2.
- Harrison, G. (2008) "Neuroeconomics: A Critical Reconsideration". *Economic & Philosophy*. Vol. 24, N° 3.
- Harrison, G. y Ross, D. (2010) "The methodologies of neuroeconomics". *Journal of Economic Methodology*. Vol 17. N° 2. pp. 185-196.

- Heilman, R., Crisan, L., Houser, D., Miclea, M. y Miu, A. (2010) “Emotion Regulation and Decision Making under Risk and Uncertainty”. *Emotion* Vol. 10, 2, pp. 257-265.
- Hubert, M. (2010) “Does neuroeconomics give new impetus to economic and consumer research?” *Journal of Economic Psychology*. (En prensa).
- Hutchison, T. (1938) *The Significance and Basic Postulates of Economic Theory*. New York: A. Kelley.
- Keynes, J. M. (1924) “Alfred Marshall”. En *Essays on Biography*. Cambridge: Cambridge University Press.
- Khaneman, D. (2009) “Remarks in Neuroeconomics”. En Glimcher, P., Camerer, C., Fehr, E. y Poldrack, R (eds.). *Neuroeconomics: Decision Making and the Brain*. Londres: Elsevier
- Kinkaid, H. y Ross, D. (2009) *The Oxford Handbook of Philosophy Economics*. Oxford: Oxford University Press.
- Koppl, R. (1991) “Retrospectives: Animal Spirits”. *Journal of Economic Perspectives*. 5, pp 203-210.
- Kuhnen, C y Knutson, B. (2005) “Neural Basis of Financial Risk Taking”. *Neuron*, 1, pp. 763-770.
- McKeil Carter, K, Meyer, J. y Huettel, S. (2010) “Functional Neuroimaging of Intertemporal Choice Models: a Review” *Journal of Neuroscience, Psychology and Economics*. Vol.3, 1, pp. 27-45.
- Mohr, P., Li, S. y Heekeren, H. (2010) “Neuroeconomics and aging: Neuromodulation of economic decision making in old age”. *Neuroscience and Biobehavioral Review*. 34, 678-688.

- Navarro, A. M. (2007) “Neuroeconomía y teoría de los juegos: implicancias metodológicas”. *Anales de la Academia Nacional de Ciencias Económicas*.
- Navarro, A. M. (2008) “Economía, biología y evolución” Conferencia Inaugural de la 43^a. Reunión Anual de la Asociación Argentina de Economía Política. www.aaep.org.ar
- Rick, S. (2010) “Losses, gains and brains: Neuroeconomics can help to answer open questions about loss aversion”. *Journal of Consumer Psychology*, Junio.
- Ross, D. “Two Styles of Neuroeconomics”. *Economics & Philosophy*. 24, pp. 473-483.
- Schwartz, H. (2010) “Does Akerlof and Shiller’s Animal Spirits provide a helpful new approach for macroeconomics?”. *Journal of Socio-Economics*. 39, pp. 151-154.
- Sobel, J. (2009) “Neuroeconomics: a Comment on Bernheim”. *American Economic Journal Microeconomics*. Vol 1. N° 2.
- Stuphorn, V. (2006) “Cardinal Utility in the Orbitofrontal Cortex?”, *Current Biology*. Vol 15, 8, pp. 591-593.
- Vercoe, M. y Zak, P. (2010) “Inductive modelling using causal studies in neuroeconomics: brain and drugs.” *Journal of Economic Methodology*. Vol 17. N° 2.

Análisis de Mercados No Convencionales: Mercados de Órganos y Repugnancia

Julio J. Elías
(Universidad del CEMA)

1. Introducción

El análisis económico brinda una serie de herramientas que permiten analizar una amplia variedad de problemas sociales y económicos desde una perspectiva tanto positiva (i.e., predictiva) como normativa (i.e., evaluativa). En este capítulo se presentan dos aplicaciones novedosas, relacionadas entre sí, y poco convencionales del enfoque económico referidas a problemas microeconómicos de relevancia tanto social como económica.

Los órganos para trasplantes humanos son bienes extremadamente valiosos y su faltante se ha transformado en uno de los problemas de políticas públicas más candentes en la mayoría de los países con programas importantes de trasplantes de órganos. En la primera parte de este capítulo se presentan los principales elementos del análisis económico de los mercados de órganos para trasplantes humanos, desarrollado por Becker y Elías (2007). En este análisis, se utilizan herramientas de la nueva economía de la mortalidad y de otras áreas del análisis económico para estimar el precio de un riñón para trasplantes y estudiar los efectos de la in-

troducción de incentivos monetarios sobre la cantidad de trasplantes, el bienestar social y otras variables de interés.

Luego se introduce el concepto de repugnancia como un límite al desarrollo de ciertas transacciones monetarias, como ser justamente la compra y venta de órganos humanos para trasplante. La eficiencia económica es el criterio que se utiliza comúnmente en el análisis económico para establecer un orden entre distintas alternativas de política. Sin embargo, en muchas situaciones parecería ser que este criterio no es el que prima. En esta segunda parte del capítulo se presentará lo que algunos dieron en llamar como el “yuck factor”, o factor repugnancia, y se examinará cómo este factor opera como restricción sobre algunas transacciones en el mercado y cuáles son sus consecuencias. Este concepto, desarrollado por Al Roth (2007), plantea que algunas transacciones, como ser la compra y venta de riñones para trasplantes o la compra y venta de niños recién nacidos, son ilegales simplemente porque un número suficiente de personas las encuentran repugnantes. Utilizando un análisis económico simple, se mostrará que lo que resulta repugnante puede llegar a depender de las circunstancias y se encuentra cercanamente asociado al costo social generado por la prohibición o regulación.

2. Introducción de Incentivos Monetarios en el Mercado de Órganos¹

Los trasplantes de órganos empezaron en 1954 con un trasplante de riñones entre hermanos gemelos realizado en el Brigham & Women’s Hospital de Boston en los Estados Unidos. Sin embargo, la cantidad de estos procedimientos recién empezó a despegar en los años setenta con el desarro-

¹ Esta sección del capítulo está basada en el trabajo de Becker y Elías, 2007.

llo de drogas inmunosupresoras, que ayudan a prevenir el rechazo del cuerpo humano a los órganos trasplantados. A partir de este progreso, el número de trasplantes de riñón, hígado, corazón y de otros órganos creció rápidamente. El resultado fue la formación de largas colas de espera para recibir órganos en la mayoría de los países con programas importantes de trasplantes de órganos.

Actualmente, en nuestro país 9.800 personas aguardan un órgano para recibirlo en trasplante. 5.500 de estas personas se encuentran en la lista de espera renal. Para lograr acceder a un trasplante renal, estos pacientes deberán sufrir una espera promedio de 5 años. La falta de riñones disponibles para trasplantes genera serios problemas en los pacientes. Mientras dura la espera, deben someterse a diálisis durante 6 horas diarias tres días a la semana. Un mayor período en diálisis afecta negativamente la tasa de éxito del trasplante y la expectativa de vida de los pacientes. Pero quizás el mayor efecto adverso de la falta de órganos es el sufrimiento y el deterioro de la calidad de vida de los pacientes mientras esperan por un órgano.

Cuando un economista observa una brecha entre demanda y oferta en algún mercado su próximo paso, normalmente, es buscar por obstáculos que puedan estar impidiendo que el mercado se equilibre. Estos obstáculos son obvios en el caso del mercado de trasplantes de órganos ya que en muy pocos países –con las excepciones de la India en los ochentas y principios de los noventas e Irán empezando en 1988– se permite el uso de incentivos monetarios para adquirir órganos. Utilizando el análisis económico, para el caso de Estados Unidos, Becker y Elías (2007) muestran que la introducción de incentivos monetarios puede aumentar la oferta de riñones para trasplantes suficientemente como para lograr eliminar las largas colas de espera en el mercado,

mediante un aumento en el costo total por trasplante de no más del 12%.

En la mayoría de los países, la principal fuente de órganos para trasplantes son los donantes fallecidos (ver Tabla 1). Debido a esto, previo al análisis de Becker y Elías, los pocos trabajos que habían discutido incentivos monetarios en estos mercados se habían concentrado en la oferta por órganos de donantes fallecidos (por ejemplo, Cohen, 1989; Becker, 1997; Kaserman and Barnett, 2002; Epstein, 1993).

De hecho, gran parte de las políticas orientadas a aumentar la oferta de órganos se han concentrado en los donantes fallecidos. Algunas de las políticas orientadas a aumentar la cantidad de donantes fallecidos incluyen la forma en que se solicita la expresión de voluntad para la donación (sistema de consentimiento presunto en lugar de consentimiento explícito), reglas otorgando prioridad en las listas de espera a aquellas personas que estaban registradas como donantes (en inglés, “*no give no take rules*”) y *Kidney Paired Donation* (ver Roth et al., 2004). Becker y Elías destacan el potencial de utilizar incentivos para obtener más donaciones vivas, aunque también consideran la posibilidad de pagar por órganos de donantes fallecidos.

La Situación del Mercado

Para evaluar el rendimiento del sistema de trasplantes de manera apropiada es conveniente advertir que la cantidad de pacientes en lista de espera en un momento dado del tiempo refleja principalmente la historia del funcionamiento sistema y no así su estado corriente, ya que la misma es la resultante de la acumulación de los faltantes de órganos de períodos pasados.

En el caso argentino el panorama luce complicado si se consideramos los 5500 pacientes que se encuentran en lista de espera renal, con un tiempo promedio de espera de 5

años. Sin embargo, debemos notar que en el año 1998 cerca de 5.200 personas se encontraban en lista de espera para recibir un riñón con un tiempo de espera promedio por encima de 11 años. Esta reducción en el tiempo promedio de espera se debió principalmente al rápido aumento en la cantidad de trasplantes que se produjo durante los últimos años, de cuatrocientos trasplantes en 1998 a mil cincuenta en el 2010. Es importante remarcar que cerca del 80% de esta expansión de la oferta está explicada por el incremento en la cantidad de donantes reales fallecidos.

Una forma de estimar la brecha entre demanda y oferta de trasplantes de órganos es comparar el número anual de trasplantes con el número anual de personas que se suma a la lista de espera en necesidad de un trasplante. Esta brecha es igual al crecimiento anual en la lista de espera más el número de personas que abandona la lista de espera por fallecimiento u otras razones. Para la Argentina, considerando únicamente el crecimiento en la lista de espera, lo que brindaría una estimación de la cota inferior de esta brecha, se obtiene que la oferta sobrepasó a la demanda en seis de los últimos diez años. Lamentablemente durante los tres últimos años la oferta de órganos fue insuficiente para satisfacer la demanda anual, lo que generó un aumento en la lista de espera renal de cerca de 800 pacientes.

Las largas colas de espera no son únicas del sistema argentino, también se encuentran en otras naciones. La tabla 1 muestra, para varios países, el número total de trasplantes renales, el número de trasplantes con donantes vivos, y el tamaño de la lista de espera por millón de población. Estados Unidos con cerca de 6 años de espera promedio es el país con la mayor cantidad de personas en lista de espera renal. A excepción de España, en todos los países de la tabla los tiempos de espera promedio superan los tres años. Existe un amplio debate acerca de las fuentes del suceso del siste-

ma español, algunos lo explican por el sistema de donante presunto, otros señalan que se debe a la efectividad de los equipos de recuperación de órganos para realizar la solicitud de donación en los hospitales o que se debe a la gran cantidad de muertes en accidentes de autos.

Incentivos Monetarios y el Análisis Económico

El sistema actual del mercado legal por trasplantes renales puede representarse por la curva de oferta SS y la curva de demanda DD en la Figura 2, en donde el eje horizontal mide el número total de trasplantes renales por unidad de tiempo, y el eje vertical el costo total de una cirugía típica de trasplante renal. En la mayoría de los países, el costo total de las cirugías de trasplantes es principalmente pagado por el gobierno o por las compañías de seguro de salud. Su predisposición a calificar personas para esta costosa cirugía aumenta a medida que el costo disminuye. Esta es la razón principal por la cual la cantidad demandada efectiva, dada por la curva DD , aumenta cuando el costo cae.

La demanda de personas de altos ingreso que se financian gran parte del costo del trasplante claramente también tiende a estar negativamente relacionada con el costo. Bajo el sistema actual, algunas personas utilizan el mercado negro en otras partes del mundo, como ser en el sur de Asia y algunos países de América del Sur, para pagar por donantes vivos o por órganos de cadáveres de manera de reducir su tiempo de espera.

Cuando los donantes o sus familias no reciben compensación por sus órganos, uno esperaría que la oferta de órganos no dependiera mucho del precio por encima del costo de la cirugía. Es por esto que la curva de oferta de trasplantes legales con órganos de cadáveres o de donantes vivos es una curva en forma de L invertida. El tramo horizontal corresponde al costo promedio del trasplante renal, y la curva de

oferta se vuelve vertical, o independiente del precio, para un número de donantes altruistas Q^0 .

En la Figura 2 se asume que el costo promedio de un trasplante renal es de us\$160.000. A ese precio, la cantidad demandada de riñones excede la cantidad ofrecida, y se forma una lista de espera o se produce un aumento en la misma. La brecha sería menor a un costo mayor porque la cantidad demandada se reduce ante un aumento en el costo. Por lo tanto, el progreso tecnológico que redujo el costo de los trasplantes de órganos y que los hizo más seguros durante los últimos quince años aumentó el exceso de demanda y la lista de espera mediante un aliento de la demanda.

La introducción de pagos monetarios a los donantes o sus familias aumentaría el costo total de los trasplantes, ya que el costo incluiría estos pagos. Pero debido a que con una mayor compensación se obtendría una mayor cantidad de órganos, esto rotaría el segmento vertical de la curva de oferta en el sentido de las agujas del reloj, de manera que la curva de oferta se movería de SS a S^*S^* . Si el gobierno subsidia o el seguro privado cubre el pago a los donantes, un mercado legal también afectaría la posición de la función de demanda redistribuyendo parte de la demanda del mercado negro al sector legal y alentando a una mayor cantidad de pacientes a aplicar para un trasplante ya que la espera se reducirá significativamente. Pero aún así, los incentivos monetarios ayudarían a cerrar la brecha entre la cantidad ofrecida y la cantidad demanda aumentando el número de órganos disponibles y mediante un aumento del costo total de la cirugía de trasplante.

De hecho, de acuerdo a este análisis los incentivos monetarios deberían cambiar la oferta de órganos de ser completamente inelástica con respecto al costo para pasar a ser altamente elástica. La razón es que el número potencial de órganos utilizables de donantes vivos y fallecidos, especial-

mente de donantes vivos, es muy grande comparado con el número de trasplantes. Pero el sistema actual bloquea prácticamente toda esta oferta potencial mediante la prohibición a cualquier forma de compensación por órganos. A continuación se muestra cómo se puede estimar lo que costaría utilizar incentivos monetarios para atraer una cantidad suficiente de donantes vivos de manera de eliminar la brecha entre oferta y demanda.

El Precio de un Riñón para Trasplante

La donación de un órgano para trasplante en vida puede llegar a afectar la calidad de vida, el riesgo de mortalidad y la habilidad para realizar actividades en el mercado y afuera del mismo por algún período de tiempo después de la cirugía. ¿Cuánto dinero se requiere para inducir a una persona a vender un órgano? Se puede estimar el precio de un órgano para un donante vivo calculando cuánto ingreso adicional o consumo de mercado requeriría una persona para estar indiferente entre vender un órgano o no. Específicamente, se puede asumir que el precio de reserva de un órgano tiene principalmente tres componentes compensadores aditivos: una compensación monetaria por el riesgo de muerte, una compensación monetaria por el tiempo perdido antes y después de la cirugía y una compensación monetaria por el riesgo de una reducción en la calidad de vida.

La compensación monetaria requerida por el vendedor de un órgano para enfrentar el riesgo de morir como resultado del trasplante, el primer componente del precio de reserva, está dado por el riesgo de morir multiplicado por la prima monetaria requerida para enfrentar un pequeño aumento en la probabilidad de morir. Este último factor, al que se lo conoce como el valor estadístico de la vida en la nueva economía de la mortalidad, ha sido estimado empíricamente en muchos estudios utilizando información proveniente de

mercados de actividades riesgosas que envuelven un “*trade-off*” entre dinero y riesgos de fatalidad. Basado en estos estudios, el valor promedio del valor estadístico de la vida para una persona joven con un ingreso anual cercano a los us\$ 35 mil en los Estados Unidos es de us\$ 5 millones (ver Viscusi y Aldy, 2003).

El riesgo para un donante de morir como resultado de un trasplante renal es menor a un 1/100 de un 1 por ciento. Asumiendo de forma conservadora que el riesgo de donar un riñón es de 0.1 por ciento y un valor estadístico de la vida de us\$ 5 millones, como en Murphy and Topel (2006), esto quiere decir que un aumento de 1/1000 en la probabilidad de morir para un donante de riñón sería valuado en us\$ 5 millones x 0.1 por ciento, o en us\$ 5.000.

La compensación monetaria por el tiempo perdido durante al recuperación, el segundo componente del precio de reserva, estará dado por los ingreso perdidos. Uno de los principales determinantes del valor estadístico de la vida es el ingreso de las personas. Como se menciona arriba, un valor de la vida de us\$ 5 millones corresponde a una persona con un ingreso anual promedio de unos us\$ 35 mil aproximadamente. Un donante vivo sufrirá una pérdida de us\$ 2.700 de ingresos como resultado de no trabajar durante 4 semanas debido a la cirugía.

La compensación monetaria requerida para enfrentar el riesgo de una reducción en la calidad de vida, el tercer componente, estará dado por el cambio esperado en la calidad de vida multiplicado por la prima monetaria requerida para enfrentar una pequeña reducción en la calidad de vida. No se posee una buena manera de valuar los efectos de donar un riñón sobre la calidad de vida. Sin embargo, algunos estudios muestran que una persona que dona uno de sus riñones en vida puede llevar una vida normal, excepto si se trata de atletas o personas que participan de actividades con mucho

contacto físico que pueden llegar a dañar el único riñón que les queda. Algunos estudios también mencionan que los donantes vivos poseen una mayor chance de desarrollar presión sanguínea alta. De forma relativamente arbitraria se agregan us\$ 7.500 para medir el componente de calidad de vida del precio (ver Becker y Elías para un mayor detalle y para un análisis de sensibilidad).

Agregando estos tres componentes se obtiene, para los Estados Unidos, un costo total esperado para los donantes de us\$ 15.200. En relación al número de personas que necesitan un riñón para trasplante, una gran oferta de donantes vivos estarán disponibles a us\$ 15.200 por donante aproximadamente.

Para verificar estos números en la Tabla 2 se presentan algunas estimaciones del precio pagado por riñones en distintos países. Como se menciona al inicio de esta sección, los pagos a donantes vivos eran permitidos en la India durante los ochentas y principios de los noventas, y los pagos continuaron siendo comunes aún después de la prohibición. También se presentan datos para Irán en donde actualmente se encuentra permitido el pago por riñones para trasplantes de donantes vivos. El precio promedio en la India (fila 3 de la tabla 2) era de us\$ 1.177 en dólares del 2005. Esta estimación fue calculada por Goyal, Mehta, Schneiderman, and Sehgal (2002) de una encuesta conducida en el 2001 entre 305 personas que vendieron un riñón en la India 6 años antes de la encuesta en promedio.

El ingreso real per cápita en la India es cerca de 1/15 del de los Estados Unidos cuando se utiliza el poder adquisitivo de compra para ajustar por diferencias en los costos de vida entre la India y los Estados Unidos. Si, por lo tanto, multiplicamos el pago en la India por 15, los datos implican que el costo de un riñón de donante vivo en los Estados Unidos sería cercano a los us\$ 17.000 dólares. Este cálculo asume

una elasticidad unitaria del valor de la vida con respecto al ingreso, algo bastante plausible basado en los estudios empíricos disponibles.

Equilibrio de Mercado

Si el costo directo de un trasplante de riñón en los Estados Unidos es cercano a los us\$ 160.000, notemos que el pago a los donantes de us\$ 15.200 tiene un efecto muy pequeño sobre el costo total de los trasplantes renales con donantes vivos, pero tiene un efecto enorme sobre la elasticidad de la oferta. Los pagos a donantes convierten a la oferta de ser inelástica en completamente elástica a un precio que es un 9.5% por encima del precio sin compensaciones monetarias para los donantes - comparar SS con S^*S^* en la figura 3.

S^*S^* corta a la función de demanda en el punto de equilibrio de mercado e^* . Con pagos a los donantes, el número de trasplantes renales aumentaría, aunque la cantidad demandada puede caer con respecto a los niveles que están generando la lista de espera actual porque el costo total será un poco mayor. A no ser que la demanda sea altamente elástica – que es bastante improbable - el aumento de un 9.5% en el costo total debido a los pagos a los donantes no tendrá un efecto negativo importante sobre la demanda efectiva de trasplantes.

La cantidad de trasplantes de equilibrio bajo un sistema con incentivos puede ser estimada utilizando este esquema. La cantidad demandada a los precios de mercado actuales consiste en la suma de la cantidad de trasplantes más la brecha entre demanda y oferta. De ese total, se debe sustraer el cambio en la demanda actual debido al cambio en el precio. Por ejemplo, si la elasticidad de la demanda es -1, un aumento de un 9.5% en el precio reducirá la cantidad demandada en 9.5%. Con una cantidad de 13.500 trasplantes rena-

les y una brecha cercana a las 8.000 personas, la cantidad demandada es cercana a los 21.500 trasplantes renales. Una caída de 9.5% en la cantidad demandada representa una caída de 2.040 trasplantes, que lleva a una nueva cantidad demandada de 19.460, todos los cuales van a ser ofrecidos. Por lo tanto después de introducir pagos a los donantes, la cantidad de trasplantes renales aumentaría de 13.500 a 19.460, o en un 44%.

Los programas de compensación para donantes vivos pueden tomar diferentes formas. En un libro reciente editado por Sally Satel del *American Enterprise Institute*, con Cronin y Elías desarrollan y evalúan un marco básico para la procuración y asignación de órganos de proveedores de órganos vivos. En el análisis se consideran dos modelos diferentes de procuración de órganos con compensación: Un sistema centralizado y uno descentralizado. El primero es un sistema regulado y centralizado en el que el gobierno o una agencia designada a tal fin actúa como la única autoridad con el poder de proveer compensación financiera y asignar los órganos para trasplantes. Así es básicamente como funciona el sistema en Irán. El segundo enfoque es un régimen en el cual el gobierno establece un conjunto de reglas y regulaciones de manera de proveer un marco legal para arreglos entre privados. Bajo estos acuerdos, el comprador y el vendedor son libres para fijar el valor de la transacción.

Lo que es más importante, en este trabajo se muestra que las principales características organizacionales de ambos sistemas de compensación serían las mismas del sistema actual de donación de órganos fallecidos y de seguimiento post trasplante de los pacientes, como es regulado en los Estados Unidos por la *Organ Procurement and Transplantation Network* (OPTN) y administrada por UNOS.

Existen ciertas características que es muy probable que sean comunes a cualquier sistema de compensación de pro-

curación de órganos. Como se menciona arriba, muchas de estas características se encuentran vigentes en el sistema actual, como ser el rol de los médicos y de los centros de trasplantes en la selección de los donantes, el cuidado post trasplante, y otros pasos del proceso de procuración de órganos. En un sistema descentralizado, el precio de un riñón será establecido a niveles que eliminarían el exceso de demanda. De acuerdo al análisis anterior, este precio sería alrededor de us\$ 15.000 en los Estados Unidos. La financiación para el pago directo a los donantes o para una remuneración en bienes puede obtenerse a partir del gran ahorro que se generaría con la reducción de la cantidad de pacientes en diálisis.

El rol de la repugnancia en el mercado

De acuerdo al análisis económico desarrollado en la sección anterior los beneficios sociales de la introducción de incentivos monetarios en el mercado de órganos para trasplantes son enormes. Teniendo en cuenta esto, una pregunta que surge de forma inmediata es por qué entonces la compra y venta de órganos humanos para trasplantes se encuentra prohibida en la mayoría de los países. En referencia al análisis económico de los trasplantes de órganos y a la posibilidad de dar lugar a un mercado de órganos, Paul Bloom, un psicólogo de Yale, expresa que “el problema no es que los economistas sean personas poco razonables, sino que son personas demoníacas. Trabajan en universo moral diferente.”

El caso de los mercados de órganos para trasplantes es un caso representativo en el cual transacciones libres y voluntarias, y aparentemente beneficiosas, son obstaculizadas por ley dentro de una sociedad debido en parte a cuestiones éticas y morales. Al Roth de la universidad de Harvard resume el arco de oposición a la introducción de compensaciones monetarias de la siguiente manera: “Legalizar la venta de riñones para trasplantes enfrenta obstáculos sustanciales,

quizás insuperables. De la misma manera que hoy no es posible venderse en servidumbre a término, algunas transacciones son ilegales porque una cantidad suficiente de personas las encuentran repugnantes.”

Es decir que algunas transacciones o mercados no se encuentran permitidas simplemente por una cuestión de gustos o preferencias en la población. Es importante destacar que a diferencia del caso estándar de externalidades en el cual la persona se ve directamente afectada porque la misma es participe en alguna medida de la transacción (ej. ruidos molestos, polución) o del mercado, en una transacción “repugnante” los participantes están dispuestos a intercambiar sin afectar de forma “material” a terceros, pero estos lo desaprueban y desean evitar que se produzca la transacción.

El área de economía del comportamiento se ha concentrado en ciertas regularidades en los gustos de las personas que generalmente no eran o no son modelados por los economistas como ser gustos fuertes por evitar pérdidas, por justicia. Los ejemplos sobre repugnancia son sobre gustos que las personas tienen respecto de las elecciones que otras personas puedan hacer – Roth (2007) muestra que esto puede tener importantes consecuencias sobre los tipos de mercados que observamos. Muchas veces la simple introducción de dinero en una transacción puede generar rechazo (ver figura 3), ya que esto puede ser visto como: Desagradable, Inapropiado, Injusto, Indigno, Poco profesional.

En la tabla 3 se presentan algunos ejemplos de transacciones que resultan repugnantes, o que alguna vez lo resultaron. Es importante notar que muchos de estos casos en los cuales la prohibición se hace o se hizo efectiva comparten ciertas características. Primero, son mercados pequeños o en los que gran mayoría de la población no piensa participar de forma directa (ej. matrimonio gay, consumo humano de carne de caballo). Segundo, parecería que la mayoría de la pobla-

ción posee un gusto no demasiado fuerte en contra de los mismos. Lo que veremos es que la combinación de estos dos elementos lleva a que en la mayoría de estos casos se imponga, o se haya impuesto en algún momento, una prohibición.

Una pregunta importante para el análisis económico es sobre los determinantes de la repugnancia y si la misma puede cambiar en el tiempo. Es decir, ¿Cómo podemos incorporar la repugnancia al análisis económico? ¿Cuál es el precio o el costo de la misma?

Para formalizar la idea de la repugnancia como un factor que puede afectar las regulaciones o costumbres, tomando el mercado de órganos como ejemplo, se empieza con la idea simple que la repugnancia de una persona hacia la compra y venta de riñones para trasplantes depende de una serie de factores.

Por ejemplo, el grado de repugnancia de una persona puede ser una función del precio esperado de un órgano. Si se espera que el precio sea muy bajo las personas pueden sentir que esto representa una forma de explotación. La repugnancia también puede depender del nivel de ingreso de los vendedores potenciales; si es muy bajo, las personas pueden sentir que es una forma de coerción hacia los sectores de menores ingresos y esto despierta repugnancia. Adicionalmente, algunos médicos envueltos en el proceso de trasplante pueden encontrar repugnante el hecho de tener que tratar con proveedores pagos en lugar de donantes altruistas.

Por otro lado, pueden existir algunos factores que operen en la dirección contraria, que afectan la repugnancia de forma negativa. Por ejemplo, algunas personas pueden encontrar repugnante que haya tantos pacientes en lista de espera, que algunos de ellos mueran mientras esperan, o que el sistema no trabaje de forma eficiente.

Para formalizar estas ideas se asume que la repugnancia de una persona i puede representarse mediante una función. Es decir

$$(1) \quad \text{Repugnancia}_i = R_i(\text{Factores})$$

Esta relación puede variar entre diferentes personas. En lo que sigue, utilizando un marco simple se analiza cómo la repugnancia afecta la decisión de la persona con respecto a permitir o prohibir la venta de órganos.

***Bienestar de las personas bajo dos regímenes:
Mercado y Prohibición***

Primero, se considera el bienestar de los individuos bajo dos regímenes diferentes. Uno es un régimen en el cual los pagos o cualquier otra forma de incentivos financieros para proveedores de riñones se encuentran prohibidos y el otro es un régimen en el cual se encuentra permitida la compra y venta de riñones para trasplante.

Nuevamente, se utiliza una función para representar el bienestar de las personas, y por simplicidad, se asume que el bienestar de los individuos es una función de tres factores únicamente: Consumo, Salud y Repugnancia. Esto es, el bienestar de un individuo i en el régimen j está dado por

$$(2) \quad \text{Bienestar Individual}_{i,j} = U(C_{i,j}, H_{i,j}, R_{i,j}) \quad j = \text{Prohibición, Mercado}$$

Es importante notar que estos tres factores pueden variar entre personas (i) y entre regímenes (j). Es decir que personas diferentes tendrán diferentes niveles de bienestar dentro de un mismo régimen y entre regímenes.

Sin embargo, para los participantes indirectos en el mercado, el nivel de consumo y su salud serán aproximadamente iguales bajo los dos regímenes. Esto es probable que sea el caso para individuos que no necesitan un trasplante, o para aquellos que no planean proveer un riñón para trasplante en caso que se abra el mercado.

De acuerdo a esta representación el bienestar de la persona i bajo los dos regímenes estará dado por las siguientes expresiones

$$(3) \quad \text{Bienestar Individual}_{\text{Prohibición},i} = U(C_{\text{Prohibición},i}, H_{\text{Prohibición},i}, R_{\text{Prohibición},i})$$

$$(4) \quad \text{Bienestar Individual}_{\text{Mercado},i} = U(C_{\text{Mercado},i}, H_{\text{Mercado},i}, R_{\text{Mercado},i})$$

La ausencia de un mercado de riñones para trasplante representa un costo en términos de bienestar para algunas personas (i.e. personas para las cuales $U(C_{\text{No Market},i}, H_{\text{No Market},i}, R_{\text{No Market},i}) < U(C_{\text{Market},i}, H_{\text{Market},i}, R_{\text{Market},i})$), mientras que para otros representa un beneficio (i.e. personas para las cuales $U(C_{\text{No Market},i}, H_{\text{No Market},i}, R_{\text{No Market},i}) > U(C_{\text{Market},i}, H_{\text{Market},i}, R_{\text{Market},i})$). Para cuantificar económicamente el gusto o el disgusto que ocasiona el mercado se formula la vieja pregunta de diferencias igualadoras de Adam Smith: ¿Cuánto consumo, o dinero, se tiene que darle a una persona para dejarla indiferente entre los dos regímenes? Llamemos a ese monto de dinero δ . Se puede calcular δ utilizando la función de bienestar del individuo, esto es

$$(5) \quad U(C_{\text{Prohibición},i}, H_{\text{Prohibición},i}, R_{\text{Prohibición},i}) = U(C_{\text{Mercado},i} + \delta, H_{\text{Mercado},i}, R_{\text{Mercado},i})$$

Es decir que δ está dado por

$$(6) \quad \delta(C_{\text{Prohibición},i}, H_{\text{Prohibición},i}, R_{\text{Prohibición},i}, C_{\text{Mercado},i}, H_{\text{Mercado},i}, R_{\text{Mercado},i}) = \delta_i$$

La heterogeneidad es un fenómeno importante en este análisis. De acuerdo a la ecuación (6), δ varía en la población de acuerdo al valor que toman los factores que afectan al bienestar de las personas en cada sistema. Por ejemplo, para una paciente que necesita un trasplante es probable que δ sea negativo, no debido a que su repugnancia hacia el mercado sea baja, sino simplemente porque su estado de salud mejorará o se espera que mejore bajo un sistema de mercado.

De acuerdo a este análisis, las opiniones de las personas pueden inferirse de la siguiente manera

Si $\delta_i < 0 \Rightarrow$ la persona está de acuerdo con un mercado de riñones para trasplantes

Si $\delta_i > 0 \Rightarrow$ la persona no está de acuerdo con un mercado de riñones para trasplantes

Si $\delta_i = 0 \Rightarrow$ la persona está indiferente

Es importante notar que el nivel de consumo y el estado de salud para la gran mayoría de la población será independiente de la existencia o no de un mercado de órganos. Como consecuencia, para gran parte de la población, δ reflejará de manera plena el valor monetario o el precio implícito de la repugnancia.

La asignación de un valor monetario a la valoración individual de la introducción de compensaciones financieras en el mercado permite abordar la pregunta acerca de cómo se determina la elección del tipo de sistema o régimen. A continuación se consideran dos criterios de elección de políticas: Principio de compensación (i.e. análisis costo beneficio) y voto mayoritario (i.e. un referéndum).

¿Sistema de Mercado?: Principio (Económico) de Compensación versus Voto Mayoritario

Para determinar si resulta deseable desde el punto de vista social un sistema basado en el mercado, lo primero que hace un economista es realizar un análisis costo-beneficio. Es decir, se evalúan las ganancias provenientes de la implementación o de la remoción de la regulación y si los perdedores pueden ser compensados por los ganadores de la implementación del sistema.

Dentro del esquema simple desarrollado en la sección anterior, se pueden calcular las ganancias netas de la remoción

de la prohibición para comprar y vender órganos agregando los δ de los diferentes individuos en la población. Si la suma de los δ es negativa, entonces los ganadores con la remoción de la prohibición pueden compensar a los perdedores, y ocurre lo contrario si la suma es positiva. De acuerdo al criterio del principio de compensación

Si $\sum \delta_i < 0 \Rightarrow$ Los ganadores pueden compensar a los perdedores \Rightarrow Legalizar el mercado.

Si $\sum \delta_i > 0 \Rightarrow$ Los ganadores no pueden compensar a los perdedores \Rightarrow Continuar con la prohibición.

Sin embargo, en la práctica parecería que el análisis costo beneficio no resulta suficiente para tomar la decisión o no es lo único que pesa para determinar el resultado, como explica Al Roth: “La venta de carne de caballo para consumo humano es ilegal en California, no porque se haya hecho un caso persuasivo de que los costos exceden a los beneficios, pero debido a que 4.670.524 votaron para hacerlo ilegal en un referéndum en el año 1998.”²

Dentro de este marco analítico, en un referéndum las personas con δ negativo votarán en contra de la prohibición, mientras que las personas con δ positivo lo harán a favor. Es decir que se puede calcular el resultado de la votación asignando $V_i = 1$ si la persona vota en contra de la prohibición (i.e. $\delta_i < 0$) y $V_i = -1$ si la persona vota a favor (i.e. $\delta_i > 0$). Entonces

Si $\sum V_i > 0 \Rightarrow$ Legalizar el mercado.

Si $\sum V_i < 0 \Rightarrow$ Prohibir el mercado.

² Elías and Roth, 2007.

Como se observa, en ambos casos el resultado depende de la distribución de los δ en la población. Sin embargo, con el primer criterio, el principio de compensación, la intensidad (i.e. la magnitud de δ) es importante mientras que con voto mayoritario sólo importa si la persona está de acuerdo o no con la prohibición (i.e. si δ es positivo o negativo), no el grado o la intensidad de su deseo.

La figura 4 muestra un ejemplo de una distribución de δ en la población. En este ejemplo, en la cola derecha de la distribución se encuentra un grupo de personas relativamente pequeño que se oponen de manera tajante a la idea de un mercado de riñones para trasplantes. En la cola izquierda de la distribución también se observa un pequeño grupo de personas pero en este caso a favor de un mercado de órganos. En el medio de la distribución se encuentra gran parte de la población, que puede sentir que hay algo equivocado o repugnante acerca de un mercado de órganos, pero sus sentimientos no son muy fuertes, en parte, porque no esperan verse envueltos de manera directa en el mercado.

En este ejemplo, un referéndum llevaría a la prohibición de la venta de órganos, mientras que un análisis costo beneficio podría indicar que existe una ganancia social de desarrollar un mercado de órganos. Es decir que se produce un conflicto entre eficiencia y “ética o moral”. En este ejemplo, la sociedad en su conjunto está dispuesta a pagar el costo económico de la prohibición.

Cambios en la repugnancia: Efectos de un aumento en la lista de espera renal

De acuerdo al análisis de repugnancia como un obstáculo para el desarrollo del mercado, una característica importante de la repugnancia es su impredecibilidad (ver Roth, 2007). Sin embargo, en el marco del análisis desarrollado en este capítulo lo que resulta repugnante es determinado por facto-

res subyacentes que pueden cambiar en el tiempo, en algunos casos de manera predecible. En particular, el costo social de la prohibición en un momento dado de tiempo opera como un factor implícito importante de la prohibición.

En el caso de los trasplantes de riñones, el beneficio potencial de un sistema basado en el mercado puede haber parecido bajo en relación a los costos de implementar un nuevo sistema cuando las restricciones legales fueron implementadas en los Estados Unidos y en la mayoría de los países, ya que el faltante de órganos no era tan severo. Actualmente la lista de espera renal se encuentra por encima de las 90,000 personas, y en los últimos 10 años ha crecido a una tasa promedio anual de 7%. Adicionalmente, el progreso tecnológico que ha vuelto los procedimientos de trasplantes de órganos más seguros y efectivos aumentó los beneficios potenciales de manera significativa.

Basándonos en las proyecciones de crecimiento de las adiciones a la lista de espera y la relativamente oferta fija de órganos disponibles para trasplantes, un sistema que depende en el altruismo exclusivamente alcanzará tiempos de espera para trasplantes de donantes fallecidos que pronto excederán la expectativa de vida de la mayoría de los pacientes que se encuentran en diálisis.

Tanto el crecimiento en la lista de espera como las mejoras en la seguridad y efectividad de los trasplantes son factores que pueden llegar a reducir la repugnancia hacia un sistema de mercado y, al mismo tiempo, aumentan el costo social de la prohibición.

De acuerdo a las ecuaciones 1 y 6, un aumento en la lista de espera podría generar un desplazamiento en la distribución de δ en la población hacia la izquierda y, por lo tanto, el voto de algunas personas puede modificarse. El signo de δ , y como consecuencia el voto, es menos probable que cambie para grupos con posiciones extremas (i.e. aquellos

con un δ en valor absoluto alto) que para participantes indirectos en el mercado (i.e. para aquellos en los que δ refleja únicamente el valor de la repugnancia).

Continuando con el ejemplo de la figura 4, un pequeño cambio en los δ puede llevar a un cambio en el resultado de la votación (ver figura 5). Recordemos que en el caso de voto mayoritario la intensidad no es relevante. Como consecuencia, el factor repugnancia y la existencia o no de un mercado son determinados de forma conjunta por un tercer factor subyacente, en este caso el tamaño de la lista de espera.

En la práctica esto puede funcionar de manera diferente, en la mayoría de los países existe un proceso político que determina el resultado final. En el ejemplo, existen dos grupos con posiciones extremas, por lo que es probable que los mismos se enfrenten de forma organizada en la arena política. Becker (1989) analiza un problema similar en su trabajo seminal sobre teoría de grupos de presión y uno de sus resultados principales es justamente que un determinante del resultado final del proceso político es el costo social de la política.

Lo que es Repugnante depende de las Circunstancias

Existen otros casos en la historia en los que prohibiciones o restricciones a ciertas transacciones fueron removidas debido al enorme costo social de mantenerlas. Por ejemplo, se podría pensar que los estadounidenses encontraban repugnante el beber alcohol principios del siglo pasado, por lo tanto decidieron prohibirlo en 1919. 14 años más tarde las personas cambiaron de idea debido a que los costos de la prohibición eran altos.

Alguna vez los seguros de vida fueron considerados repugnantes. En 1853, una editorial del New York Times sostenía que “aquel que asegura su vida o salud debe ser víctima de su propio desatino o de la picardía de otros.” Previo a la institución del seguro de vida, la viuda y sus huérfanos eran asistidos por sus vecinos y otras personas cercanas y

por grupos de ayuda mutua. En el siglo XIX, la protección financiera de las familias americanas se transformó en un bien comerciable. En la actualidad el seguro de vida es considerado una forma de altruismo.

Tomando el ejemplo de Al Roth sobre la repugnancia en California hacia el consumo humano de carne caballo; si el precio de la carne de vaca y otras carnes aumenta, es muy probable que la repugnancia hacia la carne de caballo disminuya, o que pueda llegar a desaparecer. De hecho, durante la Segunda Guerra mundial, New Jersey legalizó la venta de carne caballo debido al alto precio de la carne vacuna. Al final de la guerra, cuando se normalizó el mercado por otras carnes, el estado decidió prohibirla de nuevo.³

En el comunicado sobre el Premio Nobel de Medicina 2010 al padre de la fertilización in vitro se menciona la gran oposición que enfrentó Edwards: “Estos primeros resultados fueron prometedores, pero el Consejo de Investigación Médica decidió no financiar la continuación del proyecto. Sin embargo, una donación privada permitió que el trabajo continuara. La investigación también se convirtió en el tema de un acalorado debate ético que se inició por el propio Edwards. Varios líderes religiosos, éticos y científicos exigieron que el proyecto se detuviera, mientras que otros le dieron su apoyo.”

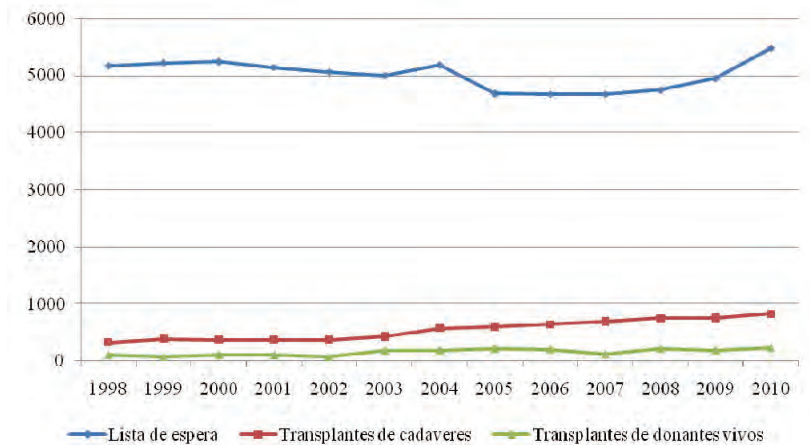
Sin embargo, lo que era repugnante hace más de treinta años, hoy representa una gran beneficio para la sociedad y es considerado por muchos un milagro de la medicina, continua el comunicado del Nobel: “Aproximadamente cuatro millones de personas hasta ahora han nacido empleando la Fecundación in vitro (FIV). Muchos de ellos son ahora adultos y algunos padres de familia se han vuelto padres de familia. Un nuevo campo de la medicina se ha convertido, con Robert Edwards quien ha dirigido el proceso hasta el final de los descubrimientos fundamentales para la terapia actual,

³ Ver http://en.wikipedia.org/wiki/Horse_meat.

el éxito de la FIV. Sus contribuciones representan un hito en el desarrollo de la medicina moderna.”

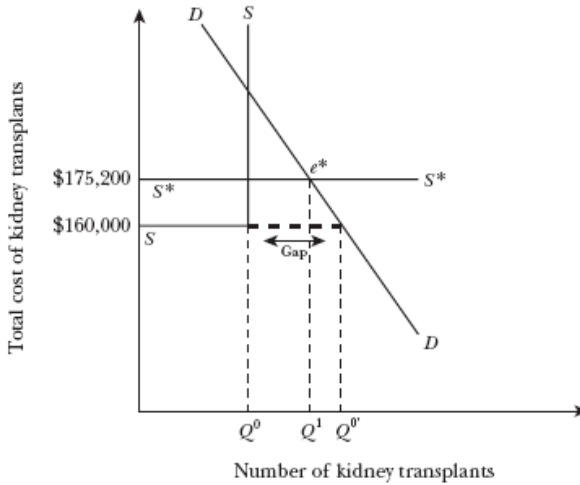
Es decir que la repugnancia responde a los costos. Cuando es barata, se observa mucha cantidad. Cuando se vuelve cara, desaparece. El análisis también nos indica que como economistas no podemos desconocer la importancia de este factor repugnancia y que el mismo puede tener importantes consecuencias sobre los tipos de mercados y transacciones que observamos.

Figura 1. Lista de Espera Renal, Trasplantes Renales con donantes vivos y de cadáveres Argentina - 1998 – 2010



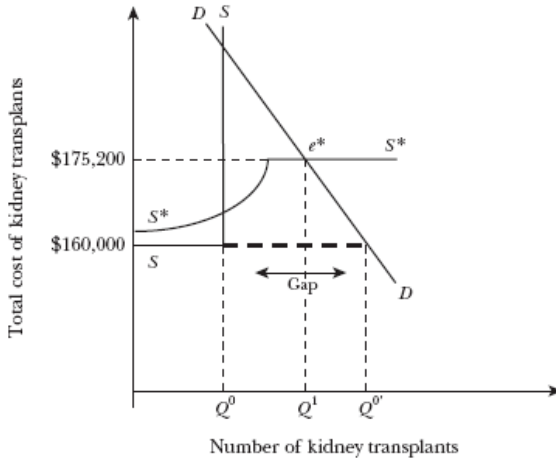
Fuente: Incucai.

Figura 2. El Mercado de Riñones para Trasplantes y Pagos para Donantes Vivos



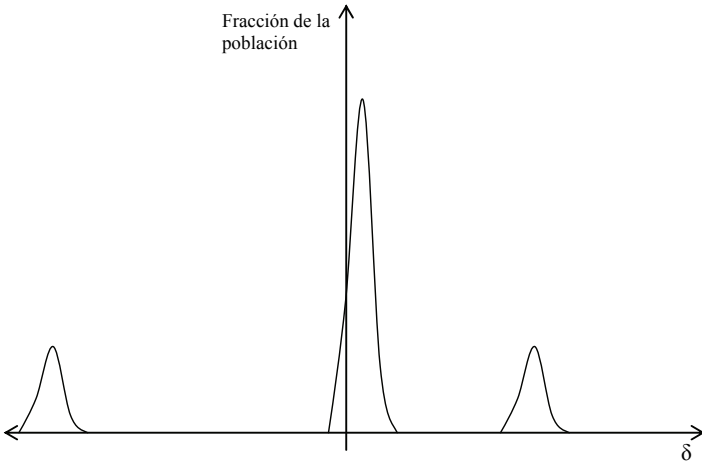
Nota: Figura tomado de Becker y Elías, 2007.

Figura 3. El Mercado de Riñones para Trasplantes con Riñones de Donantes Vivos y Fallecidos



Nota: Figura tomado de Becker y Elías, 2007.

Figura 4. Distribución de δ en la población: Un ejemplo



Nota: δ representa el monto de dinero o consumo necesario para dejar indiferente a un individuo entre el régimen de mercado y la prohibición.

Figura 5. El efecto de un aumento en la lista de espera sobre la distribución de δ en la población

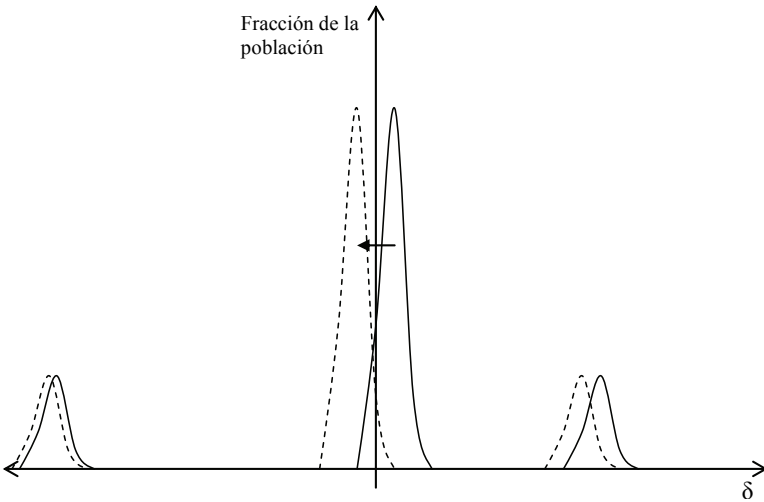


Tabla 1. Trasplantes de riñón en 2008 - 2010: Número total de trasplantes, trasplantes de donantes vivos y personas en la lista de espera. Por millón de población

	Número total de trasplantes	Trasplantes de donantes vivos	Personas en la lista de espera
Alemania (2009)	26,76	7,33	93,43
España (2009)	50,69	5,12	99,11
Reino Unido (2008 - 2009)	25,41	15,00	111,99
Estados Unidos (OPTN) (2010)	55,04	20,45	306,04
Argentina (2010)	29,29	6,51	151,08

Fuente: Cálculos propios en base a Eurotransplant, Organización Nacional de Trasplantes, UK Transplant, United Network for Organ Sharing, Incucal, y la División de Población del Departamento de Economía y Asuntos Sociales de la Secretaría de las Naciones Unidas e INDEC.

Tabla 2. Precio de un Riñón: Evidencia Internacional

País	Año	Precio del riñón en moneda local	Precio del riñón, en us\$ constantes de 2005 (h)	Costo equivalente del riñón en Estados Unidos, en us\$ constantes de 2005 (i)
Irán (a)	1997	10 millones de riales	\$1,480	\$8,536
Irán (b)	1996	3.5 millones de riales	\$518	\$2,989
India (c)	1958 - 1999		\$1,177 (d)	\$17,655
Inglaterra (e)	1990	2,500 - 3,360 libras	\$6,260 - \$8,413	\$9,077 - \$12,199
India (f)	1984 - 1988	40,000 - 50,000 rupias	\$4,281 - \$5,434	\$64,218 - \$81,510
India (g)	1997	35,000 rupias	\$1,139	\$17,078

a,b - Zargooshi (2001a).

c - Goyal, Mehta, Schneiderman, and Sehgal (2002).

d - Range \$450 - ",660. Las personas que vendieron riñones durante la década del 80, cuando era legal, recibieron \$1,603 comparado con \$975 para las personas que vendieron el riñón durante la década del 90, cuando era ilegal.

e - En 1990, un médico de Harley Street fue expulsado de la profesión luego de arreglar operaciones de trasplante en las que a cuatro turcos se les pagó entre 2,500 y 3,360 libras.

f. Salahudeen et al. (1990)

g - Evidencia de Frontline:India's National Magazine, 1997. Basada en dos caso de proveedores de riñones.

h - Calculado usando data de tipo de cambio e indice de precios del FMI.

I - Cálculos propios usando data de ingreso per cápita del World Development Indicator Database, 2002, Banco Mundial.

Nota: Tabla tomada de Becker y Elías, 2007.

Tabla 3. Ejemplos de mercados en los que algunas transacciones resultan, o que alguna vez resultaron, repugnantes

<i>Restos humanos</i>
Cadáveres para estudios anatómicos
Órganos para trasplantes cadavéricos.
Órganos para trasplantes vivos (hígados, riñones).
<i>Bebidas, comidas y drogas</i>
Carne de perro para consumo humano
Bebidas alcohólicas
Marihuana
Carne de caballo para consumo humano
<i>Riesgo</i>
Seguro de vida para adultos, niños y extraños
Juegos de azar
Especulación financiera, Intereses sobre los préstamos
<i>Polución</i>
Industrias contaminantes en países subdesarrollados.
Títulos de emisión de contaminación.

Mercados por sexo y reproducción

Madres rentadas, donación de espermias y óvulos.

Incesto

Matrimonio Gay

Mercado de Trabajo

Esclavitud, Servidumbre contratada a término

Servicio militar voluntario, soldados mercenarios

Flexibilidad laboral

Fuente: Al Roth (2007).

3. Referencias bibliográficas

Becker, Gary S. 1997. "How Uncle Sam Could Ease the Organ Shortage." *Business Week*, January 20, p.18.

Becker, Gary S. and Julio J. Elias. "Introducing Incentives in the Market for Live and Cadaveric Organ Donations," *Journal of Economic Perspectives*, Summer 2007.

Cohen, Lloyd R. 1989. "Increasing the Supply of Transplant Organs: The Virtues of a Futures Market." *George Washington Law Review*, 58(1): 1–51.

Cohen, Patricia. "Economists dissect the "Yuck" factor" *The New York Times*, January 31, 2008.

Cronin, David and Julio J. Elias. "Operational organization of a system for compensated living organ providers," *When Altruism Isn't Enough: Using Incentives to Reduce the National Kidney Shortage*, ed. Sally Satel, AEI, 2009.

Elias, Julio J. and Alvin E. Roth. "Econ One on One: A Market for Kidneys?" *The Wall Street Journal Online*, November 13, 2007.

Epstein, Richard. 1993. "Organ Transplants: Is Relying on Altruism Costing Lives?" *American Enterprise*, 4(6): 50–7.

- Goyal, Madhav, Ravindra L. Mehta, Lawrence J. Schneiderman, and Ashwini R. Sehgal. 2002. "Economic and Health Consequences of Selling a Kidney in India." *Journal of the American Medical Association*, 288(13): 1589–93.
- Kaserman, David L., and A. H. Barnett. 2002. *The US Organ Procurement System: A Prescription for Reform*. The AEI Press, Washington DC.
- Leider, S. and Alvin. E. Roth. "Kidneys for sale: Who disapproves, and why?" *American Journal of Transplantation*, 2010.
- Murphy, Kevin M. and Robert H. Topel. "The Value of Health and Longevity," *Journal of Political Economy*, 2006, vol. 114, no. 5
- Roth, Alvin E. Roth. "Repugnance as a Constraint on Markets," *Journal of Economic Perspectives*, Summer 2007.
- Roth, Alvin E. "What have we learned from market design?" Hahn Lecture, *Economic Journal*, 118 (March), 2008, 285–310.
- Roth, Alvin E., Tayfun Sonmez, and M. Utku Unver, "Kidney Exchange," *Quarterly Journal of Economics*, 119, 2, May, 2004, 457-488.
- Salahudeen, Abdulla K. et al. 1990. "High Mortality among Recipients of Bought Living- Unrelated Donor Kidneys." *The Lancet*, 336(8717): 725–8.
- Segev, D. L. and S. E. Gentry. "Kidneys for sale: Whose attitudes matters?" *American Journal of Transplantation*, 2010.
- Tabarrok, Alex. "The Meat Market," *The Wall Street Journal*, January 8, 2010.

- Viscusi, W. Kip, and Joseph E. Aldy. 2003. “The Value of a Statistical Life: A Critical Review of Markets Estimates throughout the World.” AEI–Brookings Joint Center for Regulatory Studies.
- Zargooshi, Javaad. 2001a. “Iranian Kidney Donors: Motivations and Relations with Recipients.” *The Journal of Urology*, 165(2): 386–92.
- Zargooshi, Javaad. 2001b. “Quality of Life of Iranian Kidney “Donors”.” *The Journal of Urology*, 166(5): 1790–9.

Comparación Teórica y Empírica del Índice de Malmquist y del Indicador de Luenberger. Una aplicación al mercado de seguros

Gustavo Ferro

(Instituto de Economía UADE y CONICET)

Carlos A. Romero

(Instituto de Economía UADE)

1. Introducción

El objetivo de este trabajo es presentar los avances en la medición de productividad a través de números índices e indicadores, y proveer un ejemplo centrado en el sector asegurador argentino en años recientes.

El desarrollo de indicadores e índices de productividad ha sido bastante lento. Entre el trabajo pionero de Malmquist (1953) hasta recientes artículos que estudian las propiedades de las diferentes aproximaciones para su cálculo han pasado casi sesenta años.

Hay dos aproximaciones para la medición de la eficiencia: a través de diferencias y de cocientes. Se denotan a las medidas en diferencias como “indicadores” y aquellas que se efectúan a partir de cocientes como “índices” (Diewert, 1998).

Del análisis de la literatura se desprenden las características de cada una de las medidas de productividad y las relaciones entre ellas. De tal manera que la elección entre una u otra depende de las necesidades del investigador, dadas las restricciones que le impone el problema empírico que está abordando.

El uso de índices Malmquist de productividad ha sido muy amplio en la literatura y una de sus principales ventajas es que permite descomponer el cambio de la productividad en sus elementos determinantes (Grifell-Tatjé y Lovell, 1995), en tanto su desventaja es que hay que elegir una orientación (hacia los productos o los insumos) para calcular las medidas de distancia, que en ocasiones tiene sentido (como en la orientación a los insumos si se están estudiando sectores regulados con obligación de provisión por parte del oferente), pero en otras se manifiesta como demasiado rígido o inapropiado (un mercado competitivo donde no hay restricciones a expandir la producción o reducir los insumos).

La aparición de los indicadores de Luenberger y Luenberger-Hicks-Moorsteen han permitido la utilización de modelos no-orientados, debido a que permiten comparar beneficios (considerando los componentes de ingreso y costo), o más recientemente la literatura que incluye aspectos medio-ambientales (las funciones de distancia permiten considerar tanto bienes como productos no deseados). Es decir, que se ha ganado en flexibilidad.

Aquí se presentarán formalmente estas metodologías y se realizará una comparación empírica de las mismas para ilustrar sobre sus diferencias teóricas. Para ello, se utilizará información del mercado de seguros de la Argentina, donde la elección del sector no es casual: se trata de un mercado atomizado, donde la regulación no fija cantidades, y hay también flexibilidad para modificar la mezcla de insumos. Los precios de productos e insumos pueden asumirse que se

forman bajo un patrón razonablemente competitivo. Para la aplicación empírica, las estimaciones de eficiencia relativa (tanto en forma de cocientes como con funciones de distancia) se realizarán mediante la utilización de Análisis Envolvente de Datos (DEA, por “Data Envelopment Analysis”).¹

Tras esta Introducción, en la sección 2 efectuaremos un análisis conceptual de los Indicadores e Índices de Productividad. Luego, en la sección 3, se mostrarán las relaciones entre ambos. Posteriormente, en la sección 4 se ilustra con un ejercicio aplicado al sector de seguros de Argentina. Finalmente, la sección 5 es la de conclusiones.

2. Tecnología funciones de distancia y medidas de eficiencia

Seguidamente, se expondrá conceptualmente cada método y aproximación seguida. Para ello, se presentan las definiciones de tecnología y funciones de distancia relevantes para el cálculo de los índices e indicadores y se introducen gráficamente los mismos. La estimación de la eficiencia relativa con el uso de fronteras de producción se basa en la teoría de las funciones de distancia (“Gauge Functions”). En economía, las funciones de distancia están relacionadas a la noción de “coeficientes de utilización de recursos” (Debreu, 1951) y a las medidas de eficiencia radial (Farrell, 1957).

Definiciones de tecnología y funciones de distancia

Sean $x = (x^1, \dots, x^N) \subset \mathbf{R}^N_+$ e $y = (y^1, \dots, y^M) \subset \mathbf{R}^M_+$ los vectores de insumos y productos, respectivamente, y se define la tecnología como $T_{(t)} = \{(x_t, y_t) \in \mathbf{R}^{n+p}_+ : x_t \text{ puede producir } y_t\}$. El conjunto tecnológico, $T_{(t)}$, consiste en todos los vectores factibles de insumos, x_t , y de productos, y_t , en el mo-

¹ Ferro, Lentini y Romero (2011) describen los aspectos prácticos de la medición de eficiencia relativa.

mento t que satisfacen ciertos axiomas, los cuales son suficientes para definir las funciones de distancia.²

La medida de eficiencia de Debreu-Farrell $E_{T(t)}(x_t, y_t)$ es la inversa de la función de distancia de Shepard (1953). Esta medida de eficiencia se conoce como función proporcional de distancia (FPD). Con orientación a insumos,³ esta medida representa la mínima disminución proporcional, θ , del uso de insumos para alcanzar la frontera de producción.

$$(1) \quad E_{T(t)}^i(x_t, y_t) = \min_{\theta} \{ \theta : (\theta x_t, y_t) \in T(t), \theta \geq 0 \}$$

Similarmente, la medida de eficiencia orientada a productos⁴ es la máxima expansión proporcional, θ , del nivel de producción, que sitúa la producción en la frontera con el mismo uso inicial de insumos:

$$(2) \quad E_{T(t)}^o(x_t, y_t) = \max_{\theta} \{ \theta : (x_t, \theta y_t) \in T(t), \theta \geq 1 \}$$

Luenberger (1995) generaliza las nociones previas de funciones de distancia como funciones de escasez o insuficiencia (*shortage functions*) y provee una herramienta flexible capaz de tomar en cuenta tanto contracciones de insumos como aumentos de la producción al momento de medir eficiencia. Esta “Función Direccional de Distancia” (FDD), así definida, es el dual de la función de beneficio (Luenberger, 1995; Färe y Grosskopf, 2005). La Función Proporcional de Distancia es un caso especial de la *shortage function*, que se define en el momento t como:

² La tecnología satisface los siguientes supuestos. [a]: $(0, 0) \in T(t)$, $(0, y_t) \in T(t) \Rightarrow y_t = 0$; [b]: el conjunto $A(x_t) = \{ (u_t, y_t) \in T(t) : u_t \leq x_t \}$ está acotado para todo x_t ; [c]: $T(t)$ es cerrado; [d]: para todo $(x_t, y_t) \in T(t)$, $(x_t, -y_t) \leq (u_t, -v_t) \Rightarrow (u_t, v_t) \in T(t)$; [e]: $T(t)$ es convexo.

³ Asumiendo que los productos no pueden ser alterados.

⁴ Asumiendo que los insumos no pueden ser alterados.

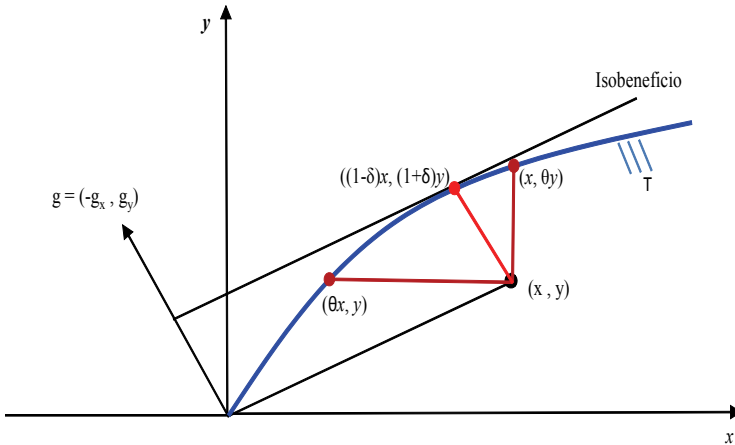
⁵ Estas medidas están relacionadas: $E_{T(t)}^o(x_t, y_t) = [E_{T(t)}^i(x_t, y_t)]^{-1}$.

$$(3) \quad D_{T(t)}(x_t, y_t) = \max_{\delta} \left\{ \delta : [(1-\delta)x_t, (1+\delta)y_t] \in T(t), \delta \geq 0 \right\}$$

donde δ es el máximo cambio para el cual el producto y_t , puede ser expandido y el insumo x_t puede ser reducido, simultáneamente, dada la tecnología $T(t)$.

La FDD permite expandir los productos en la dirección g_y y contraer los insumos en la dirección g_x . El Figura 1 ilustra esta función.

Figura 1. Función Direccional de Distancia



La tecnología T consiste en el área entre la frontera tecnológica (la curva convexa en la Figura 1) y el eje horizontal. El vector direccional $g=(g_x, g_y)$ está ubicado en el cuarto cuadrante, indicando que el producto se expande y el insumo se contrae. La FDD traslada el vector (x, y) en la dirección (g_x, g_y) hasta el límite de la tecnología, horizontalmente (contrayendo el uso del insumo x sin afectar el nivel de producto y), perpendicularmente (aumentando la producción y con el mismo uso de insumos x) o en forma oblicua (aumentando en una proporción el producto y , mientras se reduce simultáneamente en la misma proporción el uso del insumo x). Dado que (x, y) es un punto interior de

la tecnología T , el valor de la función de distancia es mayor que cero. En este caso es igual $\theta a/\theta g$, donde θg es el rayo que parte del origen hasta el punto (g_x, g_y) .

Cualquier tipo de medida, ya sea basada en cocientes o en diferencias, está fundada en la teoría de la dualidad. La función de beneficios, con su estructura aditiva, ajusta perfectamente con la FDD. En cambio, los índices de eficiencia de de Farrell (basados en costos o ingresos) están emparejados con funciones de distancia de Shepard (Färe y Grosskopf, 2005).

3. Descripción de los índices e indicadores de productividad

Primeramente, se resumirán las posibilidades de acuerdo a los caminos que ha tomado la literatura y que fueran presentados secuencialmente en la Introducción. La Tabla 1 presenta una clasificación de las medidas de productividad.

En el estudio de los índices de productividad basados en funciones distancia cabe identificar dos enfoques diferenciados. El primero arranca con Caves, Christensen y Diewert (1982), que introduce de manera teórica los denominados índices Malmquist de productividad en insumos y productos. Estos se basan en razones (cocientes) con funciones distancia orientadas a insumos, por un lado, y con funciones distancia orientadas a productos, por otro. Aunque este planteamiento tuvo una amplia difusión en la literatura teórica, la obtención empírica de índices no pudo realizarse hasta que Färe *et al.* (1994) mostraron un método de cálculo a través de técnicas de programación lineal, descomponiendo el crecimiento de la productividad así medido en dos componentes: cambio técnico y cambio de eficiencia técnica.

Tabla 1. Clasificación de medidas de productividad

Aproximaciones	Denominación aceptada	Enfoques	Literatura
----------------	-----------------------	----------	------------

a las medidas de eficiencia				
Medidas en cocientes	Índices	Malmquist	Funciones proporcionales de distancia	Caves, Christensen y Diewert (1982)
		Hicks-Moorsteen	Cociente entre Índices de Malmquist	Diewert (1992)
Medidas en diferencias	Indicadores	Luenberger	Diferencias entre Funciones direccionales de distancia	Luenberger (1995) y Chambers (1996)
		Luenberger-Hicks-Moorsteen	Diferencias entre indicadores de Luenberger	Briec y Kerstens (2004)

Fuente: Elaboración propia

El segundo enfoque referido a los índices de productividad, tiene como referencia básica a Bjurek (1996) que define una razón entre índices Malmquist en productos y en insumos. Este nuevo índice se le conoce como Hicks-Moorsteen. Esta línea ha tenido un menor seguimiento en la literatura teórica.

Por otra parte también existen dos aproximaciones para el cálculo de indicadores obtenidos por diferencias, basados en funciones de distancia. Luenberger (1995) introduce la Función Direccional de Distancia que luego es utilizada para calcular el Indicador de Productividad de Luenberger como diferencia de FDDs (Chambers, 1996). Posteriormente, Briec y Kerstens (2004) construyen el indicador de productividad de Luenberger-Hicks-Moorsteen. Estos indicadores permiten estimar los cambios de productividad con medidas de distancia que permiten reflejar cambios simultáneos en insumos y productos.

Índice de Productividad de Malmquist

El concepto de índice de productividad de Malmquist (IPM) fue introducido por Malmquist (1953), y estudiado en un contexto no paramétrico por una diversidad de autores.⁶ Se trata de un índice que representa el crecimiento de la productividad total de los factores (TFP) de una unidad de decisión (DMU).⁷ Refleja el progreso en eficiencia (tránsito hacia la frontera, si la unidad no estaba entre las inicialmente eficientes) en consonancia con el progreso de la frontera tecnológica en el tiempo (es decir, las unidades inicialmente eficientes, incrementan su productividad en el tiempo, batiendo sus propias marcas), bajo un marco de múltiples insumos y productos.

Caves, Christiansen y Diewert (1982), sugieren el uso de índices de Malmquist para realizar comparaciones entre observaciones de los niveles de producto, insumos y productividad de unidades de decisión. Podría compararse a la misma empresa en dos momentos diferentes del tiempo, por ejemplo, o bien a dos empresas diferentes en el mismo o en distintos momentos.

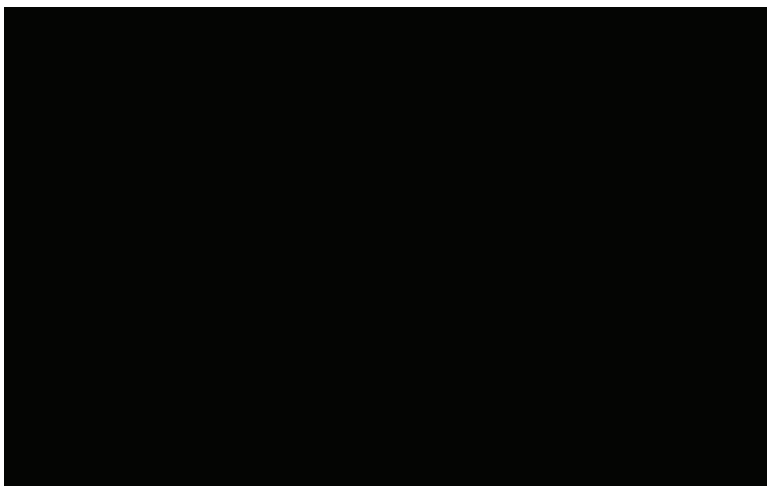
El IPM se define como el producto de dos términos: el “Movimiento de la Frontera” (también conocido como “Innovación”, “cambio técnico” o *Frontier-Shift*) y el “Cambio de eficiencia” (“Convergencia a la Frontera” o *Catch-Up*). Este último se relaciona con el grado en que una DMU me-

⁶ Färe, Grosskopf y Roos (1998) presentan una revisión de la literatura teórica y aplicada. Tone (2004) introduce conceptualmente el índice de Malmquist y presenta medidas alternativas en un marco de análisis no-paramétrico. Coelli, Rao y Battese (1998) dedica un capítulo para ilustrar como calcular este índice utilizando DEA.

⁷ Se utiliza este término porque las medidas de eficiencia y cambio de productividad no necesariamente tienen que aplicarse a empresas. Por ejemplo, es común su utilización en las áreas de salud, educación o en áreas dentro de la organización de las firmas (sucursales de bancos, departamentos de ventas regionales, entre otros).

jora su eficiencia relativo a sus comparadores, mientras que el Movimiento de la Frontera refleja los movimientos hacia mayor productividad de las DMU inicialmente eficientes, entre dos períodos de tiempo.⁸ La Figura 2 ilustra la forma de cálculo del IPM en un contexto no-paramétrico.

Figura 2. Componentes del Índice de Productividad de Malmquist



Supongamos que para una DMU los insumos y productos (x_t, y_t) observados en $t=1$ y $t=2$ son, respectivamente, los puntos $A1 = (6, 2)$ y $A2 = (4, 4)$ para el caso de un insumo y un producto. Los niveles de eficiencia relativa son 0.667 para el primer período ($B1/A1$) y 0.75 para el segundo período ($C2/A2$). Así, en $t=2$ la DMU ha mejorado su nivel de eficiencia con respecto a la frontera.⁹ Asimismo, como se observa en el figura 2, la frontera también se ha movido en el tiempo. Para medir este efecto es preciso trasladar el pun-

⁸ Un análisis de las distintas alternativas para descomponer los índices de productividad puede consultarse en Grosskopf (2003).

⁹ Dado que $C2 = (3, 4)$ y $B1 = (4, 2)$, el cambio en eficiencia es: $(C2/A2) / (B1/A1) = 1.125$, lo que indica una mejora en la eficiencia relativa.

to de referencia $B1$ a $B2$ sobre la frontera del período 2. Entonces, la convergencia hacia la frontera correspondiente a $A1$ es evaluada por $[(B1/A1) / (B2/A1)]$, y similarmente para $A2$. Para computar el efecto innovación total se toma un promedio geométrico de ambas convergencias. Finalmente, el IPM es la multiplicación de los efectos de cambio de eficiencia y cambio técnico.¹⁰

Formalmente, el índice de productividad de Malmquist, $IPM^i[(x_t, y_t), (x_{t+1}, y_{t+1})]$, orientado a insumos, se define de la siguiente manera:

$$(4) \quad IPM^i[(x_t, y_t), (x_{t+1}, y_{t+1})] = \left[\frac{E_{T(t)}^i(x_t, y_t)}{E_{T(t)}^i(x_{t+1}, y_{t+1})} \frac{E_{T(t+1)}^i(x_t, y_t)}{E_{T(t+1)}^i(x_{t+1}, y_{t+1})} \right]^{\frac{1}{2}}$$

donde,

$$E_{T(t)}^i(x_t, y_t) \mathbf{E}_{T(t)}^i(x_t, y_t), E_{T(t+1)}^i(x_t, y_t) \mathbf{E}_{T(t+1)}^i(x_t, y_t),$$

$$E_{T(t)}^i(x_{t+1}, y_{t+1}) \mathbf{E}_{T(t)}^i(x_{t+1}, y_{t+1})$$

y

$E_{T(t+1)}^i(x_{t+1}, y_{t+1}) \mathbf{E}_{T(t+1)}^i(x_{t+1}, y_{t+1})$ son las FPD (ver sección 2) necesarias para calcular el IPM^i . Esto es combinaciones de tecnología y observaciones de productos e insu-

¹⁰ Tomando $C1 = (4,5)$ y $B2 = (2,2)$, el cambio técnico es: $\left[\frac{B1}{A1} \frac{C1}{A2} \right]^{\frac{1}{2}} =$

1.826, indicando progreso técnico. En consecuencia, el IPM de este ejemplo es 2.054 que representa una mejora en la productividad total de los factores.

mos en distintos momentos del tiempo, como se ilustra en el figura 2.¹¹

La formulación DEA calcula el Indicador de productividad de Malmquist orientado a insumos¹² resolviendo el siguiente problema de optimización (bajo rendimientos constantes a escala) para cada unidad de decisión j :

$$(5) \quad E_{T(t)}^i(x_t, y_t) = \min_{\theta, \lambda} \theta \quad E_{T(t)}^i(x_t, y_t) = \min_{\theta, \lambda} \theta$$

$$Y_t \lambda \geq y_t^j$$

$$\text{Sujeto a: } X_t \lambda \leq \theta X_t^j \quad Y_t \lambda \geq y_t^j$$

$$\lambda \geq 0$$

donde θ es el nivel de eficiencia en el año t , λ es un vector $N \times I$ de ponderadores, Y_t y X_t son las matrices de productos y_t , e insumos x_t para todas las DMUs de la muestra. Los vectores de productos e insumos de la empresa j son y_t^j y x_t^j , respectivamente.

Para estimar el cambio en productividad en el tiempo, son necesarias distintas funciones de distancia orientadas a insumos y productos, incluyendo funciones de distancia del

¹¹ Adicionalmente, la descomposición para el IPM orientado a insumos es: $IPM^i = \left[\frac{E_{T(t-1)}^i(x_{t-1}, y_{t-1})}{E_{T(t)}^i(x_t, y_t)} \right] * \left[\left(\frac{E_{T(t)}^i(x_{t+1}, y_{t+1})}{E_{T(t+1)}^i(x_{t+1}, y_{t+1})} \right) * \left(\frac{E_{T(t)}^i(x_t, y_t)}{E_{T(t+1)}^i(x_t, y_t)} \right) \right]$, donde el primer término representa el Cambio en Eficiencia y el segundo término el Cambio Técnico.

¹² La formulación de los problemas de optimización con diversas orientaciones y supuestos sobre los rendimientos puede ser consultada en Cooper, Seiford y Tone (2006) y Coelli, Rao y Battese (2002).

mismo período y de períodos cruzados para cada producto y cada periodo de tiempo.¹³

Indicador de Productividad de Luenberger

El Indicador de Productividad de Luenberger (IPL), se calcula mediante una frontera tecnológica no-paramétrica para medir productividad, y no requiere una elección entre la orientación a insumos o productos (Chambers, 1996). Puesto que el Indicador de Productividad de Luenberger es consistente con ambas orientaciones, resulta ser una generalización de Índice de Productividad de Malmquist y superior al mismo por su flexibilidad (Luenberger, 1995; Färe y Grosskopf, 2005; Boussemart et al., 2003).

Formalmente, el Indicador de productividad de Luenberger, $IPL[(x_t, y_t), (x_{t+1}, y_{t+1}); g_b, g_{t+1}]$, se define de la siguiente manera:

$$(6) \quad IPL[(x_t, y_t), (x_{t+1}, y_{t+1})] = \frac{1}{2} \left[\begin{aligned} & (D_{T(t)}(x_t, y_t, g_t) - D_{T(t+1)}(x_{t+1}, y_{t+1}, g_{t+1})) \\ & + \\ & (D_{T(t+1)}(x_t, y_t, g_t) - D_{T(t+1)}(x_{t+1}, y_{t+1}, g_{t+1})) \end{aligned} \right]$$

Al igual que el IPM, el Indicador de Productividad de Luenberger también puede ser expresado como la suma los componentes: cambio técnico y cambio en eficiencia.¹⁴

¹³ Así, para el caso de las observaciones en el período inicial, t, con respecto a la tecnología en el periodo siguiente el problema a resolver es: $E_{T(t+1)}^i(x_t, y_t) = \min_{\theta, \lambda} \theta$ sujeto a: $Y_{t+1} \lambda \geq y_t^j$; $X_{t+1} \lambda \leq \theta x_t^j$; $\lambda \geq 0$, y para la comparación entre la DMU en el nuevo período (t+1) con respecto a la tecnología en el período inicial el problema es: $E_{T(t)}^i(x_{t+1}, y_{t+1}) = \min_{\theta, \lambda} \theta$

sujeto a: $Y_t \lambda \geq y_{t+1}^j$; $X_t \lambda \leq \theta x_{t+1}^j$; $\lambda \geq 0$.

¹⁴

$IPL[D_{T(t)}(x_t, y_t) - D_{T(t+1)}(x_{t+1}, y_{t+1})] + \frac{1}{2} [(D_{T(t-1)}(x_{t+1}, y_{t+1}) - D_{T(t+1)}(x_{t+1}, y_{t+1})) + (D_{T(t+1)}(x_t, y_t) - D_{T(t)}(x_t, y_t))]$

donde la primera diferencia entre corchetes representa el Cambio en Eficiencia y el segundo término, que es la media aritmética de dos diferencias, representa el Cambio Técnico.

La formulación DEA calcula el Indicador de productividad de Luenberger resolviendo el siguiente problema de optimización, bajo rendimientos constantes a escala (Chambers, Färe y Grosskopf, 1996):

$$(7) \quad D_{T(t)}(x_t, y_t, g_t) = \max_{\theta, \lambda} \delta$$
$$Y_t \lambda \geq (1 + \delta) y_t^j$$

sujeto a:

$$X_t \lambda \leq (1 - \delta) x_t^j$$
$$\lambda \geq 0$$

donde δ es el nivel de eficiencia en el año t , λ es un vector $N \times I$ de ponderadores, Y_t y X_t son las matrices de productos y_t e insumos x_t para todas las DMUs de la muestra. Los vectores de productos e insumos de la empresa j son y_t^j y x_t^j respectivamente.

Para estimar el cambio en el tiempo de la productividad, son necesarias distintas funciones proporcionales de distancia, incluyendo tanto el mismo periodo como una mezcla de periodos, para cada producto e insumo. Para la función de distancia de “periodo-cruzado,” tenemos dos años, t y $t+1$, por ejemplo. Así, $D_{T(t)}(x_{t+1}, y_{t+1}, g_{t+1})$ es el valor de la función de distancia proporcional para el vector insumo-producto del periodo $t+1$ y tecnología en el periodo t .¹⁵

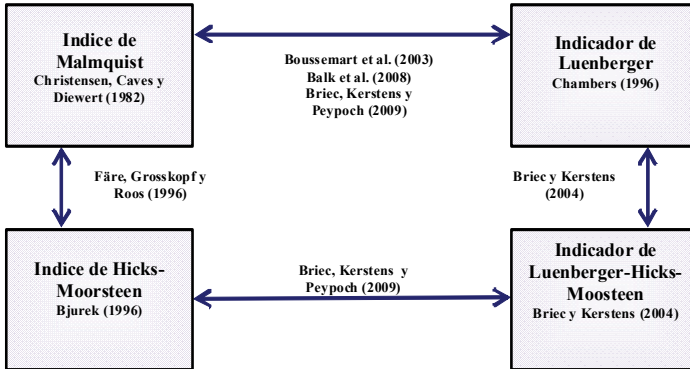
Relaciones entre los Índices y los Indicadores

Las propiedades de cada uno de estos indicadores e índices han sido estudiadas por Boussemart et al. (2003). Asimismo, Färe, Grosskopf y Roos (1996), Balk et al. (2008),

¹⁵ Siguiendo un procedimiento similar al presentado para el IPM (nota al pie 12)

Briec y Kerstens (2004), y Briec, Kerstens y Peypoch (2009) han formalizado las relaciones aproximadas y exactas entre cada una de las medidas de productividad. La Figura 3 presenta el aporte de cada uno de los artículos a la literatura.

Figura 3. Contribuciones de la literatura



Fuente: Briec, Kerstens y Peypoch (2009)

En este trabajo nos concentramos en la relación existente entre el índice de productividad de Malmquist y el indicador de productividad de Luenberger. Boussemart et al. (2003) encuentran una relación aproximada entre estas dos medidas del cambio de la productividad. Muestran que el logaritmo del Índice de Productividad de Malmquist orientado a insumos es dos veces una aproximación lineal del Indicador de Productividad de Luenberger (en valor absoluto).

$$(8) \quad \log \{IPM^i [(x_t, y_t), (x_{t+1}, y_{t+1})]\} = 2 IPL [(x_t, y_t), (x_{t+1}, y_{t+1})]$$

Dado que la medida orientada a insumos implica sólo reducciones en los factores de producción, sobreestima la in-

eficiencia técnica. Como consecuencia, se espera que también sobrestime los cambios de productividad.¹⁶

Por otra parte, Balk et al. (2008) y Briec, Kerstens y Peypoch (2009) establecen la relación exacta entre los IPM e IPL, pero solamente para aquellas DMUs que están sobre la frontera en ambos periodos para los que se evalúa el cambio de productividad.¹⁷

4. Aplicación al Sector Seguros de la Argentina

Antes de los 1990s el sector tuvo un crecimiento acotado, especialmente en el segmento de vida por la alta inflación, combinada con la ausencia de una unidad de cuenta indexada para definir los contratos financieros. En los 1990s la posibilidad de denominar contratos en dólares, así como la reforma de las pensiones en 1994 y del régimen de accidentes de trabajo en 1996, permitió un desarrollo de los ramos personales. La crisis financiera de 2002 tuvo varias influencias sobre el sector: desaparecieron los contratos en dólares, la caída del PBI motivó la retracción de la demanda, que se comenzó a recuperar junto con la economía desde 2003, y posteriores cambios institucionales en pensiones, donde se eliminó el sistema de capitalización, motivaron una caída continua del segmento de vida, a la par que la inflación retornaba en años recientes. La devaluación de la moneda implicó una fuerte reducción en las primas expresadas en dóla-

¹⁶ Similarmente, el logaritmo del Índice de Productividad de Malmquist orientado a productos es dos veces el Indicador de Productividad de Luenberger (en valor absoluto).

¹⁷ Briec, Kerstens y Peypoch (2009) demuestran que el indicador de productividad de Luenberger orientado a productos se relaciona con el índice de productividad de Malmquist de acuerdo con la siguiente expresión:
$$IPL^o = \frac{1}{2} \sum_{\tau=0,1} (2\tau-1) (IPM_{\tau}^o)^{2\tau-1}.$$

res, que recién hacia 2008 llegaron a los niveles pre-crisis, de unos us\$ 7000 millones.

La penetración del sector (producción medida en primas en relación al PBI) está en la media regional, aproximadamente en 2,5% del PBI, baja en relación a algunos países vecinos (como Chile) y mucho más en comparación a países desarrollados (como España). La densidad (los dólares per cápita de primas), estaba en us\$ 184 antes de la crisis y la devaluación, y recién había recuperado esos niveles en 2008 (Swiss Re, 2009).

Base de Datos

La base de datos en uso, confeccionada por los autores a partir de información oficial centralizada por la Superintendencia de Seguros de la Nación), consta de cinco observaciones anuales para 71 compañías actuantes en el período 2004 a 2008,¹⁸ sobre cinco variables indicativas de producción, todas expresadas en unidades monetarias, y cuatro de insumos también denominadas en unidades monetarias. Se está abarcando aproximadamente la mitad del mercado en términos de primas netas totales. Las observaciones que no pudieron ser incluidas en la muestra contenían faltantes de datos en variables específicas, o bien inconsistencias. El criterio de inclusión fue completitud y consistencia de los datos sobre las variables utilizadas.

Los ingresos de la industria se clasifican en “técnicos”, como ser primas vendidas para cobertura de siniestros –en ramos personales y no personales-, y reaseguros tomados por las compañías y financieros, por ejemplo por la colocación en instrumentos de deuda de las reservas que se man-

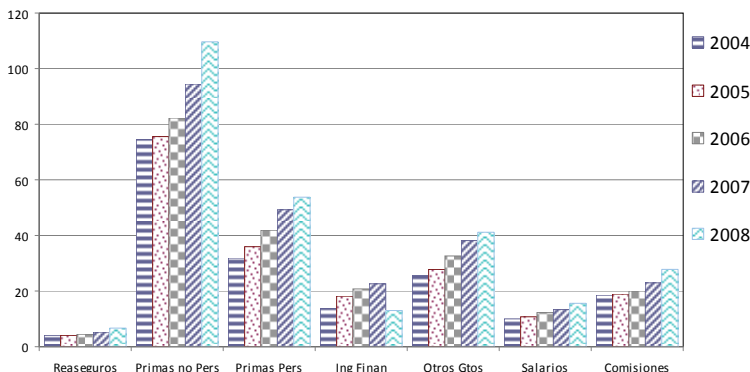
¹⁸ Los ejercicios contables de las aseguradoras, por disposición regulatoria, cierran a fin de junio de cada año. De ese modo, 2004 debe ser interpretado como el período que corre entre el 1 de julio de 2003 y el 30 de junio de 2004, y así para los demás.

tienen en cartera (Baltensperger et al., 2008). La proporción de primas por ramos no personales ha crecido desde un 60% a un 80% recientemente. Los ingresos financieros han sido muy volátiles y los reaseguros tomados tienen un porcentaje bajo y estable con relación a las primas. La industria en su conjunto cede un 20% de sus primas a reaseguradores extranjeros. Aquí sólo se contempla la cesión de primas entre compañías nacionales (reaseguros tomados por las compañías de otras aseguradoras).

La estructura de costos, en tanto, refleja el proceso productivo de una industria que vende directamente algunos seguros, principalmente de ramos personales, en tanto los no personales se colocan mayormente por el canal de intermediarios (productores asesores de seguro). Aproximadamente un tercio de los costos totales son comisiones a intermediarios, un poco más del 18% salarios y algo menos del 50% restante son otros gastos. Dicha estructura de costos se mantuvo en el tiempo.

La Figura 4 permite efectuarse una idea inicial de los cambios en productividad del sector, al examinarse como crecieron los rubros de ingreso y los de costos en el período. Las primas en ramos no personales crecieron más rápidamente que en ramos personales, en tanto los reaseguros lo hicieron más lentamente y los ingresos financieros tuvieron una evolución diferente. Dentro de los componentes de costos, las comisiones y los otros gastos crecieron más rápidamente que los salarios. Los componentes de ingreso tuvieron en conjunto un crecimiento más veloz que los de costos.

Figura 4. Evolución promedio de ingresos y costos de las aseguradoras de Argentina (muestra de 71 empresas). Período 2004-2008 (en millones de pesos de 2008)



Fuente: Elaboración propia sobre datos de la Superintendencia de Seguros de la Nación

En la Tabla 2 se presenta la estadística descriptiva de los promedios para todo el período de la base de datos utilizada para realizar las estimaciones de productividad.

Tabla 2. Estadísticas descriptivas de las variables empleadas en las estimaciones.

Período 2004-2008	Reaseguros tomados	Primas ramos no personales	Primas ramos personales	Ingresos Financ.	Otros gastos	Salarios	Comisiones
Mínimo	0	0	0	0	1516	237	0
Máximo	56822	924688	364664	241873	418933	194340	252441
Promedio	4788	87343	42553	17559	32993	12374	21654
Desvío Estándar	9555	159614	63827	30088	53578	20850	33040
Dispersión	2	1,8	1,5	1,7	1,6	1,7	1,5

Fuente: Elaboración propia sobre datos de la Superintendencia de Seguros de la Nación.

La Tabla 3 permite caracterizar mejor la muestra para cada uno de los períodos, mostrando la distribución relativa de los ingresos por primas totales (incluyendo reaseguros tomados) de las compañías.

Se observa que el tamaño de las empresas aumenta entre 2004 y 2008. En alrededor de un 54% de las empresas de la muestra tenían primas menores a 50 millones de Pesos de 2008, en cambio en 2008, la participación de empresas en este rango cae a 34%. El tamaño relativo de las empresas de la muestra es dispar (niveles de desvío estándar superiores a la media en todos los años). Sin embargo, se observa que la dispersión disminuye punta a punta (de 1.53 a 1.40), lo que significa que la distribución del tamaño de las empresas disminuyó.

Tabla 3. Frecuencia relativa (%) de los ingresos por primas totales (primas personales y no personales más reaseguros tomados) de empresas de seguros de la Argentina. Período 2004-2008 (en millones de Pesos de 2008).

Rango	2004	2005	2006	2007	2008
menor a 20	19.7	18.3	12.7	12.7	11.3
20.01 – 50	33.8	32.4	32.4	28.2	22.5
50.01 – 100	16.9	18.3	22.5	21.1	25.4
100.01 – 200	14.1	14.1	15.5	16.9	15.5
200.01 – 500	12.7	14.1	11.3	14.1	15.5
mayor a 500	2.8	2.8	5.6	7.0	9.9
Total empresas	71	71	71	71	71
Promedio	110	116	129	149	170
Desvío estándar	169	172	181	206	238
Dispersión	1.53	1.48	1.40	1.39	1.40

Fuente: Elaboración propia

Resultados

Dado que no se dispone de datos físicos para producción, sino solamente para insumos (personal y sucursales) se optó por variantes donde tanto los productos como los insumos están expresados en términos monetarios. En este sector no hay direccionalidad en cuanto a productos o insumos en el logro de la optimización empresarial: las decisiones implíci-

tas en los cambios de productividad son tanto de utilización de insumos como de expansión o contracción de la producción. Por la naturaleza competitiva del sector se asume que los precios de productos y costos están igualados dentro del mercado argentino de seguros.

Los productos considerados son primas de seguros agrupadas en personales y no personales. Los insumos son salarios, comisiones y otros costos. Se asume rendimientos constantes a escala y se consideraron modelos radiales y en diferencias orientados a insumos y productos y un modelo en diferencias no orientado.¹⁹

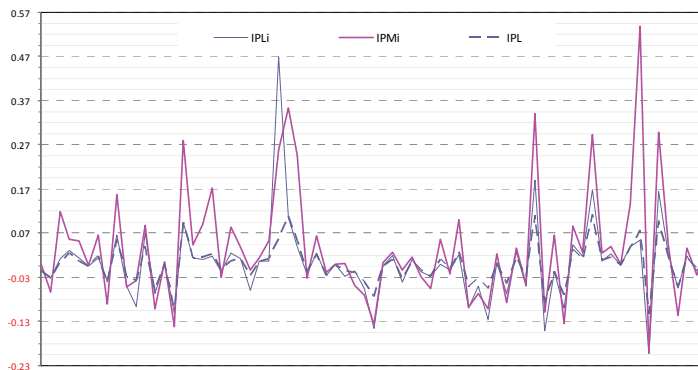
Debido a que el índice de productividad de Malmquist orientado a productos es la inversa del orientado a insumos, las comparaciones se realizan tomando en cuenta este último con respecto a los indicadores de Luenberger (no orientado) y Luenberger orientado a insumos (Figura 5).

Se observa, tal cual lo menciona la literatura, que el índice de Malmquist sobreestima el crecimiento de la productividad. En la Figura 5, este índice aparece con niveles más altos en la mayoría de los casos. Sólo para una empresa en particular, el índice de Luenberger (insumos) es más alto. El índice de Malmquist también sobrestima la caída de la productividad, aunque aquí hay cuatro casos donde el IPL orientado a insumos da una caída mayor. Sin embargo, el objetivo es comparar con el IPL, que es la medida adecuada en un ambiente donde las unidades de decisión pueden tanto afectar la producción como su demanda de insumos.

Concentrándonos en el índice de Malmquist y el indicador de Luenberger, de acuerdo con la literatura existe una relación entre los niveles de productividad estimados por ambos métodos.

¹⁹ Eling y Luhn (2008) y Cummins y Weiss (1998) realizan un análisis de las aplicaciones de modelos de frontera aplicados al sector de seguros.

Figura 5. Comparación de tasas de crecimiento estimadas con el Índice de Malmquist (insumos) y los indicadores de Luenberger (insumos y no orientado), por empresa. Promedio del período 2004-2008.



Fuente: Elaboración propia.

Nota: IPL: Indicador de Productividad de Luenberger (no orientado); IPLi: Indicador de Productividad de Luenberger orientado a insumos; IPMi: Indicador de Productividad de Malmquist orientado a insumos

Por un lado, existe una relación exacta (Balk et al., 2008; Briec, Kerstens y Peypoch, 2009) para aquellas empresas que son eficientes en ambos periodos para los que se calcula el cambio de productividad. Se ha comprobado que las estimaciones realizadas cumplen con las relaciones exactas especificadas.

Por otra parte existe una relación aproximada para toda la muestra de empresas, tanto eficientes como no eficientes (Boussemart et al. 2003). La Tabla 4 presenta la comparación de las diferencias entre ambas medidas de productividad, comparando el crecimiento anual (2005 representa el cambio entre 2005 y 2004, y así sucesivamente). Se toma como medida el cociente entre ambos términos de la ecuación anterior, y se espera un valor cercano a uno. Se observa

que no se puede tomar el resultado aproximativo como una regla. En promedio la aproximación muestra que las medidas de Malmquist están entre 15% a 22% por encima de la medida de Luenberger, con una variabilidad por empresa relativamente baja (dispersión entre 0.20 y 0.30).

Tabla 4. Frecuencia relativa (%) de las diferencias entre el Índice de Malmquist y el Indicador de Luenberger. Período 2004-2008.

Rango	2005	2006	2007	2008
hasta 1%	12.7	16.9	19.7	21.1
entre 1% y 5%	23.9	14.1	22.5	25.4
entre 5% y 10%	21.1	23.9	14.1	15.5
entre 10% y 20%	12.7	16.9	19.7	16.9
entre 20% y 50%	19.7	18.3	15.5	11.3
Más de 50%	9.9	9.9	8.5	9.9
Total	100.0	100.0	100.0	100.0
Promedio	1.22	1.21	1.16	1.15
Desvío estándar	0.37	0.41	0.23	0.25
Dispersión	0.31	0.34	0.20	0.22

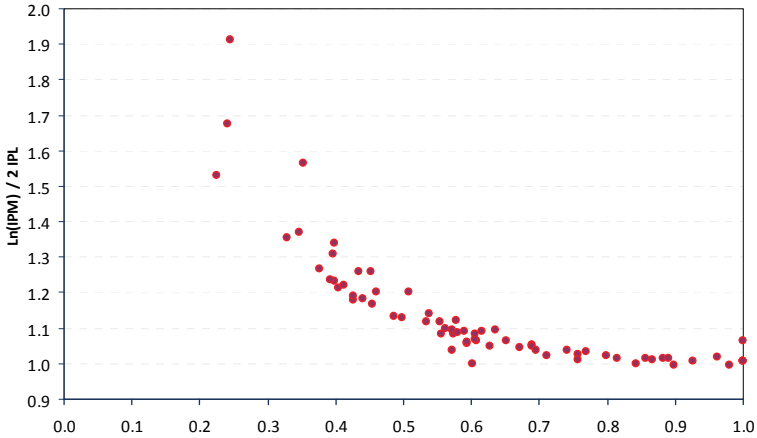
Fuente: Elaboración propia

La frecuencia relativa expone adicionalmente, que esta aproximación no puede ser utilizada directamente a nivel de cada unidad de decisión. Si bien no menos de 55% de los casos se encuentran dentro de un error del 10% (año 2006), se observa que los errores mayores al 20% llegan a casi el 30% de los casos (en el año 2005).

Este resultado indica que se requiere mayor información para poder comparar resultados de productividad obtenidos por intermedio del Índice de Malmquist y del Indicador de Luenberger. A continuación se estudia si las diferencias se pueden explicar por los niveles relativos de eficiencia.

La Figura 6 presenta la relación entre las diferencias calculadas de acuerdo a Boussemart et al. (2003) y los niveles promedio de eficiencia, calculados de forma radial orientados a insumos.²⁰

**Figura 6. Diferencias entre IPM e IPL y nivel de eficiencia.
Promedio del período 2004-2008.**



Fuente: elaboración propia.

Se observa claramente que a menores niveles de eficiencia, los errores se magnifican. En otras palabras, a medida que el nivel de eficiencia se acerca a la frontera, las diferencias de estimación se desvanecen. Sin duda este resultado es sólo conjetural dado que se basa en una muestra particular de empresas de seguros de la argentina. Sin embargo, está en línea con la literatura que se dedicó a buscar las relaciones exactas a partir de suponer eficiencia de las unidades productivas o a partir de aproximaciones que tienden a uno cuando la variabilidad de la muestra es baja.

²⁰ Si bien la figura muestra resultados para el promedio de los cuatro años considerados para cada DMU, los mismos se mantiene en términos cualitativos cuando se toma cada período individualmente.

5. Conclusiones

A la forma tradicional de medición de cambios en productividad, Índices de Malmquist, expresados como razones, se han agregado en los últimos años nuevas medidas, Indicadores de Luenberger, formadas como diferencias. Balk et al. (2008) y Briec, Kerstens y Peypoch (2009) han hallado las relaciones matemáticas entre ambos, guardando proporcionalidad las medidas para empresas eficientes (las que están sobre la frontera) y también se han encontrado relaciones aproximadas entre las diferentes medidas de cambio de productividad para todas las empresas de la muestra (Boussemart et al., 2003).

Los Índices de Malmquist resultan de utilidad en el análisis económico cuando el problema a resolver tiene una orientación clara: esto es, si las unidades de decisión trabajan ajustando sus insumos por alguna restricción al cambio en la producción (por ejemplo en sectores regulados de infraestructura) o bien si la orientación es al producto (por alguna rigidez en el cambio en la mezcla de insumos).

En tanto, en determinadas industrias la orientación a insumos o a productos no es una buena descripción de las alternativas que cuentan las unidades de decisión para sus objetivos de maximizar beneficios o minimizar costos. Los Indicadores de Luenberger ofrecen medidas no orientadas. A la vez, permiten descomponer las variaciones totales en la productividad en cambio técnico o en cambio en eficiencia.

Resulta importante contar una relación entre las diferentes medidas de cálculo de la productividad, para poder hacer correctas comparaciones de diferentes mercados diferentes momentos del tiempo. En este trabajo se muestra que los resultados obtenidos hasta el momento, si bien importantes, todavía son parciales, y se contribuye conjeturando que muestras donde la variabilidad de la eficiencia es alta pueden afectar las comparaciones aproximadas. Adicionalmente

se muestra que a menor eficiencia es más difícil realizar comparaciones entre indicadores. Una extensión de este trabajo requeriría la utilización de *bootstrapping* para estudiar con mayor profundidad la conjetura que surge del este trabajo, variando el tamaño y la composición de la muestra.

6. Referencias bibliográficas

- Balk, B., R. Färe, S. Grosskopf y D. Margaritis, 2008. “Exact relations between Luenberger productivity indicators and Malmquist productivity indexes,” *Economic Theory* 35, pp. 187-190.
- Baltensperger E., P. Buomberger, A. Iuppa, B. Keller y A. Wicki, 2008. “Regulation and intervention in the insurance industry – fundamental issues.” *The Geneva Reports*, Geneva Association.
- Bjurek, H., 1996. “The Malmquist total factor productivity index,” *Scandinavian Journal of Economics* 98, pp. 303-313.
- Boussemart, J-P., W. Briec, K. Kerstens y J-C. Poutineau, 2003. “Luenberger and Malmquist productivity indices: Theoretical comparisons and empirical illustration,” *Bulletin of Economic Research* 55, pp. 391-405.
- Briec, W. y K. Kerstens, 2004. “A Luenberger-Hicks-Moorsteen Productivity indicator: Its relation to the Hicks-Moorsteen productivity index and the Luenberger productivity indicator,” *Economic Theory* 23, pp. 925-939.
- Briec, W., K. Kerstens y N. Peypoch, 2009. “Exact Relations Between four Definitions of Productivity Indices and Indicators,” Document de travail du LEM 2009-08, abril.
- Caves, D., L. Christensen y W. Diewert, 1982. “The economic theory of index numbers and the measurement of inputs, outputs and productivity,” *Econometrica* 50, pp. 1393-1414.

- Chambers R., 1996. "A new look at exact input, output, and productivity measurement," Working Paper 96-05, Department of Agricultural and Resource Economics, University of Maryland.
- Coelli, T., D. Prasada-Rao y G. Battese, 1998. *An Introduction to Efficiency and Productivity Analysis*, Kluwer Academic Publishers, Boston.
- Cooper, W., L. Seiford y K. Tone, 2006. *Introduction to Data Envelopment Analysis and Its Uses*, Springer, New York.
- Cummins, J. y M. Weiss, 1998. "Analyzing Firm Performance in the Insurance Industry Using Frontier Efficiency Methods," Financial Institutions Center 98-22, The Wharton School, University of Pennsylvania.
- Debreu, G., 1951. "The Coefficient of Resource Utilization," *Econometrica* 19, pp. 273-292
- Diewert, W., 1998. "Index Number Theory Using Differences Rather than Ratios," Department of Economics, UBC, Discussion paper 98-10.
- Eling, M., M. Luhnen, 2008. "Frontier efficiency methodologies to measure performance in the insurance industry: Overview and new empirical evidence," *Working papers on risk management and insurance* 56.
- Färe, R., S. Grosskopf, B. Lindgren y P. Roos, 1994. *Productivity Development in Swedish Hospitals: A Malmquist Output Index Approach*, Kluwer Academic Publisher, Boston.
- Färe, R., S. Grosskopf y P. Roos, 1996. "On two definitions of productivity." *Economic Letters* 53, pp. 269-274.
- Färe, R, S. Grosskopf y P. Roos, 1998. "Malmquist productivity indexes: a survey of theory and practice," en R.

- Färe, S. Grosskopf y R. Russell (eds): *Index Numbers: Essays in Honour of Sten Malmquist*, Kluwer Academic Publishers, Norwell, MA.
- Färe, R, S. Grosskopf y G. Whittaker, 2004. “Distance Functions With Applications to DEA,” en W. Cooper, L. Seiford y J. Zhu (eds): *Handbook on Data Envelopment Analysis*, Kluwer Academic Publishers, Norwell, MA.
- Färe, R y S. Grosskopf, 2005. *New Directions: Efficiency and Productivity*, Springer, New York.
- Farrell, M., 1957. “The Measurement of Productivity Efficiency,” *Journal of the Royal Statistical Society* 120, pp. 253-281.
- Ferro, G, E. Lentini y C. Romero, 2011. “Eficiencia y su medición en prestadores de servicios de agua potable y saneamiento. Una guía útil para gerentes y reguladores”. *Documento de Proyecto LC/W.385 CEPAL/GTZ*. Santiago de Chile, febrero.
- Grifell-Tatjé, E. y C. Lovell, 1995. “A Note on the Malmquist Productivity Index,” *Economics Letters* 47, pp. 169-75.
- Grosskopf, S., 2003. “Some Remarks on Productivity and Its Decompositions,” *Journal of Productivity Analysis* 20:3, pp. 459-474
- Luenberger, D., 1995. *Microeconomic theory*, New York: McGraw-Hill Book Co.
- Malmquist, S., 1953. “Index Numbers and Indifference Surfaces,” *Trabajos de Estadística* 4, pp. 209-242.
- Shepard, R., 1953. *Cost and Production Functions*, Princeton University Press, Princeton

Swiss Reinsurance Company Ltd., 2009. “Insurance in Latin America: Focus on Personal Lines,” Swiss Re Insights, November.

Tone, K., 2004. “Malmquist Productivity Index: Efficiency Change Over Time,” en W. Cooper, L. Seiford y J. Zhu (eds): *Handbook on Data Envelopment Analysis*, Kluwer Academic Publishers, Norwell, MA.

Producción y Oferta. Teoría y Aplicaciones en Economía Agropecuaria

Daniel Lema

(IES-INTA y Universidad del CEMA)

1. Introducción

La tecnología de producción es un tema que ocupa una parte importante del análisis económico teórico y empírico por el lado de la oferta. La economía aplicada al sector agropecuario tiene una larga tradición en este tópico que puede remontarse a los estudios pioneros de Earl Heady en la Universidad de Iowa. En las investigaciones aplicadas se han enfatizado los aspectos vinculados con la oferta de productos, las demandas de factores, el cambio tecnológico, la distribución del ingreso, las relaciones entre precios de factores y productos, los retornos a escala, el tamaño y distribución de empresas y la acumulación de capital. En muchos casos, las conclusiones que se derivan de las aplicaciones empíricas dependen del orden de magnitud de los parámetros involucrados en la definición y estimación de la tecnología. Por lo tanto, el análisis conceptual y los aspectos de medición son de crucial importancia.

Este capítulo presenta una revisión de los aspectos más relevantes de la teoría de la producción y sus aplicaciones, con especial referencia a cuestiones empíricas relevantes

que se han desarrollado en muchos casos alrededor de estudios en economía agropecuaria. La primera parte desarrolla los temas de teoría y estimación de funciones de producción, luego se presenta la aproximación dual a la tecnología como avance conceptual y empírico de relevancia. Se analizan también los aspectos de la tecnología de producción y su vinculación con las estimaciones de productividad y cambio tecnológico. Finalmente, se introduce el tema de producción bajo incertidumbre y se presenta el enfoque de producción estocástica y los avances recientes en el análisis de producción bajo estados contingentes.

2. El Enfoque de la Función de Producción

En principio, tanto desde el punto de vista teórico como aplicado a temas de economía agropecuaria, existen dos grandes períodos en el estudio de funciones de producción: antes y después de la difusión del enfoque de dualidad. El cambio en la aproximación al problema se produce hacia principios de los años setenta y se produce por los avances metodológicos en la teoría y su vinculación con la estimación empírica.

El trabajo pionero a nivel microeconómico para estimar funciones de producción corresponde a Heady (1946) y tiene como antecedentes una nota metodológica de Tintner (1944) y una aplicación de Tintner y Brownlee (1944)¹ publicadas en el *Journal of Farm Economics*.

Estos trabajos microeconómicos estaban influenciados por el estudio agregado de Cobb y Douglas (1928), cuyo objetivo central había sido poner a prueba la teoría de la retribución marginal de los factores y estimar las participaciones relativas del capital y el trabajo en el producto.

¹ A su vez, estos trabajos estaban influenciados por el ya reconocido estudio de Cobb y Douglas (1928).

Las motivaciones de los estudios de Heady eran principalmente estimar las relaciones técnicas entre insumos y productos y experimentar con nuevas formas para la función de producción.

Estos estudios utilizaban una muestra aleatoria de empresas agropecuarias clasificadas por tamaño y se incluían como insumos la tierra, el trabajo, equipos de capital, stock de animales, alimentación del ganado y gastos operativos. Un punto interesante es que ya estos trabajos incorporaban explícitamente la dificultad de medición de un aspecto que aún hoy continúa siendo relevante en las aplicaciones empíricas, tal como lo es el factor empresarial o “management” de las empresas.

Otro tema relevante era el de la agregación y la estimación de los parámetros de interés de una función de producción a partir de la variabilidad observada entre empresas. Es decir, la estimación se refiere a variaciones de uso de insumos y factores entre empresas y no dentro de las empresas. Lo que sugiere que, desde el punto de vista conceptual, podrían existir múltiples funciones de producción dada la diversidad de uso de factores y niveles de producción. Lo cual implica que las estimaciones de funciones de producción de largo plazo serían “híbridos” que representan muy aproximadamente las condiciones tecnológicas preponderantes.

Los estudios que siguieron a esta primera generación, a lo largo de más de veinte años, centraron su atención en la estimación de las contribuciones factoriales, las economías de escala, el rol del cambio tecnológico, las elasticidades de producción y de oferta, la definición práctica de los insumos y el tratamiento del factor empresarial. Una característica importante de este grupo de estudios empíricos es la falta de robustez de los resultados de las estimaciones. En general, los resultados son condicionales al conjunto de información utilizado y muchas veces diferentes muestreos de los mis-

mos datos producían resultados distintos en términos de elasticidades factoriales. Heady y Dillon (1961, capítulo 17) reportan una importante cantidad de estudios de diferentes países basados en micro-datos de productores agrícolas. Las elasticidades promedio y sus coeficientes de variación son:

Tabla1. Elasticidades Insumo-Producto

	Tierra	Trabajo	Otros Servicios
Elasticidad	0.38	0.21	0.39
C. V.	0.58	0.80	0.59

Fuente: elaboración propia en base a Heady y Dillon (1961)

Se observa que en promedio la suma de las elasticidades es cercana a la unidad (retornos constantes a escala). Sin embargo, la magnitud del coeficiente de variación indica una gran dispersión en los resultados obtenidos, lo cual implica que deben tomarse con cierto cuidado al momento de interpretar y generalizar los coeficientes estimados a partir de estas funciones de producción.

Asimismo, en el texto de Heady y Dillon (1961) pueden encontrarse los primeros resultados de formas funcionales flexibles hallados a través de la expansión de series de Taylor como aproximaciones polinomiales a funciones de producción. Estos autores postularon una forma polinomial logarítmica de segundo grado que adicionaba términos cuadráticos y cruzados a la clásica función Cobb-Douglas. Diez años más tarde, Christensen, Jorgenson y Lau (1971) formalizaron y fundamentaron las propiedades de esta forma funcional que se conoce con el nombre de “translogarítmica”. Heady y Dillon también presentaron los primeros resultados de estimaciones por mínimos cuadrados ordinarios (MCO) de una transformación cuadrática que incluía como caso particular la función de producción lineal generalizada que luego fue formalmente introducida por Diewert (1971).

Las estimaciones directas de funciones de producción por parte de Heady et al. utilizando MCO estaban basadas en experimentos controlados, es decir, las variables del lado derecho de la ecuación de regresión (cantidades de insumos o factores) estaban por definición no correlacionadas con el término de error. Sin embargo, en el caso no experimental (datos de empresas), incorporar cantidades de insumos como variables explicativas implica un grave problema de simultaneidad.

Es decir, dado que las cantidades de insumos utilizadas son resultado de una decisión económica de optimización, las mismas no pueden considerarse exógenas y en consecuencia los estimadores de MCO de los parámetros de la función de producción serán sesgados.

Una forma de solucionar este problema de simultaneidad fue propuesta originalmente por Klein (1953). El método propuesto fue estimar las proporciones factoriales respecto del producto, calculadas a partir del supuesto del cumplimiento de las condiciones de primer orden para maximización de beneficios. Por supuesto, esto implica imponer las condiciones de competencia y en consecuencia las mismas no pueden ser testeadas. Los estimadores basados en el enfoque de dualidad se originan en estos trabajos e implícitamente comparten esta característica.

3. Un modelo simple de producción

El problema de optimización y el enfoque de la dualidad entre producción y costos es microeconómico en esencia. En consecuencia, sus aplicaciones y las implicancias empíricas son relevantes para la interpretación de datos a nivel de empresas, más que a datos agregados a nivel sectorial o de mercados. La discusión de los principales aspectos de la relación entre teoría y aplicación empírica del modelo de función de producción pueden presentarse utilizando un

modelo simple basado en una función de tipo Cobb-Douglas con un solo insumo (Mundlak 2000):

$$(1) \quad Y = AX^\beta e^{m_0 + \mu_0},$$

donde m_0 es un efecto específico de la empresa (por ej. *management*²) conocido por la empresa pero no por el investigador (puede pensarse como información privada o no revelada por la muestra) y μ es un término aleatorio desconocido por la empresa al momento de tomar las decisiones de producción.

La esperanza condicional del producto, dado el nivel de insumo, de la empresa i es:

$$(2) \quad Y_i^e \equiv E(Y|X_i) = AX_i^\beta e^{m_0}.$$

Suponemos que el precio es conocido y que la empresa elige el nivel de insumo, X , de forma tal de maximizar el beneficio esperado:

$$(3) \quad \max_{X_i} \pi^e(X|w, p, i) = PY_i^e - X_i,$$

donde P es el precio real del producto expresado en unidades de insumo.³ La condición de primer orden puede expresarse incorporando del lado derecho de la ecuación dos términos aleatorios adicionales que captan el efecto estocástico en el proceso de decisión económica:

² Otras fuentes de efectos específicos en empresas agropecuarias pueden ser las diferencias en calidad de la tierra, clima, etc. No obstante, el énfasis usualmente es puesto en el factor de manejo empresarial. Este efecto específico de empresa no sólo es observado en estimaciones realizadas con datos a nivel de empresas agropecuarias sino también en análisis de sección cruzada en empresas industriales o de servicios. En consecuencia, el efecto de variables ambientales para empresas agropecuarias puede no ser la principal razón de los efectos idiosincrásicos.

³ O de manera equivalente, el precio del insumo se supone $W=1$

$$(4) \quad \frac{dY^e}{dX} = \beta AX^{\beta-1} = \frac{1}{P} e^{m_1 + \mu_1},$$

donde m_1 es conocido para la empresa, pero no para el investigador y μ_1 es un componente transitorio. El término m_1 , por ejemplo puede representar la formación de expectativas y μ_1 un shock transitorio sobre los precios. Para obtener el sistema de producción (demanda de insumo y oferta de producto) utilizamos las ecuaciones (2) y (4). En cada una transformamos las variables en logaritmos y las medimos en desvíos con respecto a su media. Incorporando un subíndice temporal:

$$(5) \quad y_{it} - x_{it}\beta = m_{0it} + \mu_{0it},$$

$$(6) \quad y_{it} - x_{it} = -p_{it} + m_{1it} + \mu_{1it}$$

Si suponemos que los precios son exógenos, la forma reducida para x será:

$$(7) \quad x_{it} = (1 - \beta)^{-1}(p_{it} - \mu_{1it} - m_{1it} + m_{0it} + \mu_{0it}),$$

Los componentes de error se suponen independientes e idénticamente distribuidos (iid):

$$(8) \quad \mu_{jit} \sim (0, \sigma_{jj}) ; \quad m_j \sim (\mu_j, \tau_{jj}); \quad j= 0, 1$$

El supuesto de independencia de μ en el tiempo implica que no existe correlación serial. En consecuencia, es posible pensar a la perturbación μ como un reflejo de variables tales como condiciones climáticas, ambientales o de mercado que la empresa no puede anticipar. Pueden existir shocks que afectan a las todas las empresas en un período determinado, tales como sequías o inundaciones y por lo tanto generar correlación en μ entre las observaciones individuales. Estos shocks generan un efecto temporal que puede ser controlado de la misma manera que un efecto individual por empresa.

Las observaciones acerca de la endogeneidad de los insumos se vinculan con el resultado presentado por la ecuación (7). Esta ecuación muestra que el uso del insumo es función del efecto individual, m_{0i} , el cual también es parte del shock sobre la función de producción y por lo tanto el insumo no es estrictamente exógeno. El sesgo causado por esta endogeneidad es lo que contribuye a la falta de robustez y diversidad de resultados en las estimaciones, en particular a las diferencias observadas en estimaciones intra y entre empresas (series de tiempo vs. corte transversal). Asimismo, cuando los coeficientes estimados con sesgo se utilizan para testear eficiencia en el uso de recursos, comparando productividades marginales con precios reales, pueden inducir conclusiones erróneas acerca de la eficiencia en la asignación o uso de los recursos en las empresas.

Dado que en estas condiciones, la estimación por MCO no sólo es sesgada sino inconsistente, se han propuesto varias soluciones al problema de endogeneidad.

Si la muestra consiste en datos de panel, la transformación de las variables en desvíos con respecto a la media de cada empresa permite que el efecto individual por firma desaparezca de la ecuación (7). Si denominamos a la media muestral a lo largo del tiempo x_i , entonces la ecuación (7) resulta en:

$$(9) \quad x_{it} - x_i = -c(p_{it} - p_i + \mu_{1it} - \mu_{1i} + \mu_{0it} - \mu_{0i})),$$

donde el efecto individual no está presente. El estimador obtenido en este caso es el “*within*” y está basado en las variaciones dentro de cada firma.

Un enfoque alternativo es usar el precio como una variable instrumental para estimar la ecuación (5). Esta es básicamente la aproximación para solucionar el problema por el método de dualidad. Sin embargo, el estimador obtenido con esta metodología puede ser menos eficiente, dado que no

utiliza toda la información disponible. Para describir el problema intuitivamente, de la ecuación (7) se puede inferir que la variabilidad del insumo proviene de los siguientes componentes: p_{it} , μ_{1it} , μ_{0it} , m_{1it} y m_{0it} . El último término es el que causa el sesgo y debe ser eliminado para la estimación, mientras que los anteriores proveen la variabilidad necesaria para realizar la estimación. En consecuencia, el método más eficiente sería incorporar los primeros componentes como variables instrumentales en una estimación. Sin embargo, esto no puede realizarse ya que solamente la variable p es observada. El estimador *within* utiliza la variabilidad intra firma de p y μ_i como instrumentos, mientras que el estimador dual utiliza como instrumentos la variabilidad total de p , pero no utiliza la información de μ_i que puede ser relevante. Un punto importante a considerar desde el punto de vista empírico es que cualquier variabilidad del insumo, independientemente de si es o no consistente con las condiciones de primer orden de maximización de beneficios, genera puntos en la función de producción y por lo tanto contribuye con su identificación.

Asimismo, el uso de los precios como instrumentos está sujeto además a algunas limitaciones. Si la muestra consiste micro datos de empresas competitivas, la variabilidad entre empresas de los precios será nula (o muy pequeña). Si la muestra consiste en datos de mercado, en lugar de micro datos, entonces los precios no son necesariamente exógenos y por lo tanto no pueden ser utilizados como variables instrumentales. En la aplicación empírica lo más frecuente es encontrar resultados directos de la estimación *within* y también dual, esta última a través de las participaciones factoriales utilizando precios como variables exógenas.

Otro tema importante en la investigación empírica es la determinación de los retornos a escala. Si suponemos retornos constantes a escala, en el caso de la función de un solo

insumo, el coeficiente β es ciertamente igual a uno y entonces no hay nada para estimar. El problema es relevante en el caso más realista de múltiples insumos.

El caso presentado puede generalizarse para k insumos, y en este caso el modelo consiste en la ecuación (5) donde x y β serán k vectores y k ecuaciones equivalentes a (6). En este sistema de condiciones de primer orden, si se realiza la diferencia entre las ecuaciones para dos insumos cualquiera, por ejemplo 1 y 2, en la misma quedan eliminadas las variables m_0 y μ_0 :

$$(10) \quad x_2 - x_1 = w_2 - w_1 + \mu_2 - \mu_1 + m_2 - m_1.$$

Entonces, $x_2 - x_1$ puede servir como variable instrumental, y además contiene toda la información relevante relativa a los dos insumos. Existen $k-1$ de estos instrumentos disponibles para la estimación y en consecuencia hace falta un instrumento adicional para completar la estimación del sistema. Imponer el supuesto de retornos constantes escala es una forma de solucionar el problema. En este caso, una función de tipo Cobb-Douglas donde las variables se dividen por uno de los insumos estaría libre del problema de simultaneidad y esta es una solución utilizada frecuentemente.

4. El Enfoque de Dualidad

Saliendo del caso Cobb-Douglas, las funciones de producción cuadráticas con múltiples insumos contienen una gran cantidad de variables correlacionadas entre sí, lo que hace que las estimaciones econométricas directas sean altamente sensibles al problema de la multicolinealidad. Para evitar este problema el procedimiento usual es estimar los parámetros de la función de producción a partir de las participaciones factoriales, con o sin las restricciones impuestas por la función de producción. La idea implícita es que las variaciones observadas en las proporciones factoriales en la

muestra pueden ser atribuidas a diferencias en los ratios de insumos, o de manera equivalente, a diferentes puntos sobre la función de producción. En la literatura, una de las especificaciones más usuales para la función de producción es la translogarítmica, la cual estimada directamente no provee resultados satisfactorios debido a la gran cantidad de parámetros involucrados y a los problemas de co-linealidad. En su lugar es frecuente la estimación de funciones de beneficios o costos derivadas de un proceso de optimización donde la tecnología subyacente se asume de tipo translogarítmica. En este cambio de función objetivo vía teorema de la envolvente y haciendo uso de las propiedades de dualidad, las proporciones factoriales a estimar se convierten en funciones de precios en lugar de funciones de cantidades. Un punto importante es que la idea de dualidad implica que cada punto de la función de producción se corresponde de manera unívoca con un vector de precios relativos. Sin embargo, lo contrario no es necesariamente así, salvo que se realicen una serie de supuestos restrictivos sobre las características de la función de producción, lo cual es usualmente impuesto en las estimaciones econométricas (Chambers 1988). Una vez impuestas estas restricciones, las variaciones en los precios pueden interpretarse como variaciones en las cantidades de insumos y también como causa de estas variaciones.

Bajo este enfoque, la tecnología de producción se resume en funciones de beneficios, costos o ingresos. La función de beneficios es expresada en términos de precios de productos y factores, la de costos en precios de factores y nivel de producto y la de ingresos en términos de precios de producto e insumos. En análisis de series de tiempo, en general, se incluye alguna medida de cambios en la tecnología, usualmente alguna tendencia temporal lineal o cuadrática que permite aproximar la tasa de cambio en la productividad. También

las funciones de beneficios o costos permiten incluir algunos insumos fijos y pueden ser entonces caracterizadas como restringidas o de corto plazo. De la misma forma, las funciones de beneficios o ingresos pueden ser restringidas por el lado del producto, dando lugar a estimaciones, por ejemplo, bajo cupos de producción (Fulginiti y Perrin 1993).

La dualidad se generalizó en el análisis teórico hacia fines de los años sesenta y las aplicaciones empíricas durante los años setenta y ochenta se hicieron frecuentes en la literatura (Mc Fadden 1978, Jorgenson 1986). La ventaja del enfoque, desde el punto de vista empírico, se vincula con el hecho de que para las empresas competitivas los precios, a diferencia de las cantidades, son exógenos y por lo tanto cuando se utilizan como variables explicativas no causan el sesgo de simultaneidad presente en las estimaciones directas de funciones de producción.

También, desde el punto de vista conceptual, la dualidad ofrece una forma atractiva y simple para caracterizar la estructura de producción. De manera simplificada, la dualidad significa que siguiendo la lógica de la optimización (maximización de beneficios o minimización de costos) es posible moverse desde una función de producción hacia una función dual o de comportamiento (oferta de producto y demanda de factores) y retornar a la función original (Diewert 1974). Entonces, lo que nos sugiere este enfoque es que conociendo la función de producción es posible conocer las funciones de comportamiento y viceversa. Sin embargo, esto es relativamente simple sólo si tanto la tecnología como las funciones duales tienen formas funcionales cerradas que permitan la estimación empírica de los parámetros relevantes. Ejemplos de este tipo de funciones son la Cobb-Douglas y la CES.

El problema empírico es más sutil cuando esta “*autodualidad*” no puede aplicarse directamente, como en el caso de formas funcionales complejas tales como las cuadráticas.

Por ejemplo, si se plantea una función de beneficios o costos dual, es posible preguntarse cuál puede ser la productividad marginal de un insumo o cómo es afectada esta productividad marginal por el ratio de utilización de los insumos. Dado que las condiciones de competencia son impuestas en la formulación de las funciones duales, la respuesta a la primera pregunta es relativamente simple y está implícita en la lógica dual, la productividad marginal debe siempre igualar el precio real del factor. Por otra parte, la dependencia de la productividad marginal del insumo con respecto al uso de otros factores es más compleja y difícil de responder, excepto en el caso de formas funcionales cerradas que permiten la mencionada “autodualidad”.

Por otra parte, la mayor virtud del enfoque dual desde el punto de vista empírico, está dada por la facilidad de obtener las soluciones para la estimación de los parámetros de las funciones de comportamiento y realizar estática comparativa. Así, para ser estimable y poder realizar estática comparativa, una función de producción de un solo producto y m insumos (o la correspondiente función dual) debe contener al menos $(m+1) \cdot (m+2)/2$ parámetros (Chambers 1988). Una función cuadrática que cumple con las condiciones de simetría, como por ejemplo una translogarítmica, tiene exactamente esta cantidad de parámetros y es considerada una función flexible en el sentido de que puede proveer una aproximación de segundo orden⁴ a una función de producción no conocida. Como los insumos tienden a moverse de manera conjunta, es difícil estimar de manera directa la función, entonces el procedimiento usual es estimar las propor-

⁴ Los parámetros en cuestión son primeras y segundas derivadas. Un punto importante en la estimación es que los valores serán dependientes de las combinaciones observadas de insumos y productos. En consecuencia, en el caso de mucha variabilidad en la muestra una aproximación por coeficientes fijos puede no ser muy precisa.

ciones factoriales con los datos disponibles. En este sentido, la solución práctica para la estimación dual es una extensión del estimador propuesto inicialmente por Klein (1953).

Para que las funciones duales describan el sistema de producción de manera consistente y permitan la estática comparativa deben mantener ciertas propiedades que pueden ser testeadas empíricamente. Las más directas son las propiedades de monotonicidad y convexidad (o concavidad).

Cuando la estimación trata de recuperar las demandas de factores u oferta de productos, la monotonicidad impone signos sobre las primeras derivadas de las funciones duales (oferta de producto, demanda de insumos), mientras que la convexidad implica condiciones sobre las segundas derivadas o más exactamente sobre el signo de la matriz Hessiana. Si estas condiciones no se cumplen en la estimación, el sistema es inconsistente con la maximización de beneficios supuesta inicialmente. Además de estas propiedades, la forma dual permite testear varias hipótesis acerca de la estructura de producción tales como la separabilidad, la homotecia y las propiedades del cambio tecnológico. Una característica importante de los estudios empíricos es la diversidad de resultados obtenidos y en muchos casos la dificultad de alcanzar en las estimaciones las condiciones de regularidad requeridas. Los estudios varían también en la forma funcional utilizada, el tipo de técnica de estimación y en el énfasis en cuanto a las preguntas que se proponen responder. Una síntesis de estudios y aplicaciones empíricas utilizando estos enfoques puede verse en Shumway (1995). Debe notarse que si bien la dualidad es una teoría microeconómica en su concepción, muchos de los estudios empíricos utilizan datos macroeconómicos o agregados en series temporales.

5. El Análisis de Funciones de Costos

Una de las aproximaciones posibles a la estructura de producción es a través de la función de costos dual. Se define en principio la función de costos restringida:

$$(11) \quad C(w, k, y, T) = \min_v [wv : y = F(v, k, T)],$$

donde v es un vector de insumos variables con precios w , k es un vector de insumos restringidos o fijos que se supone no tienen costo de oportunidad, y es un vector de productos y finalmente T es un índice de tecnología.

Por el lema de Shephard obtenemos las participaciones factoriales:

$$(12) \quad \frac{\partial \ln C(w, k, y, T)}{\partial \ln w_j} \equiv S_j(w, k, y, T)$$

Para el análisis aplicado se imponen usualmente diversas restricciones. Muchos de los estudios suponen que todos los insumos son variables, en cuyo caso k no es parte del problema. Asumiendo esta simplificación, podemos plantear las principales propiedades de interés para el análisis empírico.

Homotecia:

$$(13) \quad C(w, y, T) = \phi(y)C(w, T).$$

Lo que implica independencia de las proporciones factoriales del nivel de producción:

$$(14) \quad S_j(w, y, T) = S_j(w, T).$$

Cambio tecnológico neutral:

$$(15) \quad C(w, y, T) = A(t)C(w, y);$$

Entonces:

$$(16) \quad S_j(w, y, T) = S_j(w, y)$$

La homotecia y neutralidad conjuntamente implican entonces:

$$(17) \quad S_j(w, y, T) = S_j(w)$$

La función de costos es generalmente expresada como una función cuadrática o como una transformación logarítmica en las variables. Esto resulta en la función translog, que produce ecuaciones de participación factorial lineales en las mismas variables. En general, la tecnología es representada simplemente como una tendencia lineal. El análisis empírico se concentra en la estimación de las proporciones factoriales imponiendo usualmente las condiciones señaladas y la cuestión central en los estudios tiende a explicar el comportamiento de las proporciones factoriales en respuesta a cambios en precios y a analizar la tendencia de las participaciones factoriales (cambio tecnológico).

Existen dos limitaciones conceptuales importantes que pueden mencionarse respecto del enfoque de la función de costos dual. Primero, la función de costos se deriva para un agente tomador de precios, y esto no es aplicable a datos agregados o de mercado donde los precios se determinan por acción de oferta y demanda. La demanda de factores es derivada de la función de costos y por lo tanto es afectada por los shocks que afectan a la función de costos. Estos shocks se trasladan entonces a los precios de los factores. Entonces puede ocurrir que los precios no sean exógenos. Esta limitación se aplica también a los estudios que utilizan datos de mercado (incluyendo aquellos basados en funciones de beneficio). Este es un punto importante en particular para los estudios en agricultura ya que en algunos mercados,

como el de trabajo, el sector puede no ser tomador de precios, y en otros como el de la tierra definitivamente no lo es.

En segundo lugar, la función de costos es derivada de forma condicional a un nivel de producto y esto puede erróneamente llevar a pensar que el producto es exógeno. En general no hay razón para creer que el costo marginal, y en consecuencia el producto, sean independientes de los shocks sobre la función de costos. Esta segunda limitación no es compartida por el enfoque de la función de beneficios que tiene mayor flexibilidad en su aplicación.

6. El Análisis de Funciones de Beneficio

La función de beneficios es también una forma conveniente de representar una tecnología y es utilizada frecuentemente en el análisis empírico por sus posibilidades de aplicación directa.

La función puede definirse por:

$$(18) \quad \pi(p, w, k, T) = \max_{y, v} [py - vw : y, x \in T],$$

donde y es un vector de productos; x es un vector de J insumos de los cuales v son variables y k son fijos; T es la tecnología disponible; p es un vector de precios de productos y w es un vector de precios de factores.

Haciendo uso del lema de Hotelling podemos escribir la oferta de producto y la demanda de factores:

$$(19) \quad \frac{\partial \pi}{\partial p_i} = y_i(p, w, k, T).$$

$$(20) \quad \frac{\partial \pi}{\partial w_j} = -v_j(p, w, k, T).$$

Las ecuaciones de oferta de producto y demanda de insumos pueden también expresarse en términos de participa-

ciones factoriales en el beneficio. Tal como en la función de costos la función de beneficios se expresa usualmente como una función cuadrática en una transformación logarítmica de las variables, lo que resulta en expresiones lineales de las participaciones factoriales en las variables de interés.

Existe una importante literatura metodológica y de aplicación empírica de este enfoque, con estudios pioneros dedicados a la estimación de elasticidades y análisis de políticas en el sector agropecuario. En una de las primeras aplicaciones empíricas agregadas, López (1984) estimó una función de beneficios generalizada de Leontieff para la agricultura de Canadá utilizando datos de sección cruzada. Antle (1984) utiliza una translog de beneficios para estimar funciones de oferta de producto, demanda de insumos y cambio tecnológico para la agricultura de los Estados Unidos de Norteamérica entre 1910 y 1978. Un trabajo muy relevante en esta línea se refiere al sector agropecuario argentino y analiza los efectos de largo plazo de las intervenciones de política. En este estudio, Fulginiti y Perrin (1990) estiman una función de beneficios agregada, multi-producto y multi-insumo con una especificación translogarítmica. Las elasticidades de oferta de producto y demanda de insumos estimadas son utilizadas por estos autores para calcular los efectos de las políticas de intervención en el sector agropecuario entre 1940 y 1980 demostrando importantes impactos negativos en términos de producción, productividad y retornos factoriales.

7. Funciones de Producción, Heterogeneidad Tecnológica y Productividad

En general, las medidas de productividad están basadas en la comparación de cambios en el producto agregado con cambios en un índice agregado de insumos. Asimismo, el enfoque usual para estimar cambio tecnológico supone la

existencia de una tecnología única en cada momento del tiempo. De esta manera, la estimación mediante datos de series temporales permitiría calcular la tasa y las características del cambio tecnológico.

Formalmente, si escribimos la función de producción como:

$$(21) \quad Y(t) = F[A_1(t)X_1(t), \dots, A_k(t)X_k(t), t],$$

donde A_k son funciones que aumentan los factores o también índices de calidad.

Diferenciando logarítmicamente la función anterior y utilizando la notación $d \ln x / dt = \hat{x}$, obtenemos:

$$(22) \quad \hat{Y}(t) = [\omega_1(t)(\hat{A}_1(t) + \hat{X}_1(t) + \dots + \omega_k(t)(\hat{A}_k(t) + \hat{X}_k(t))] + \tau(t)$$

$$(23) \quad \hat{Y}(t) = [\text{Indice de insumos}] + \tau(t),$$

donde los ω son ponderadores y τ es el cambio relativo en la productividad total de los factores o el “residuo” que aproxima el cambio tecnológico.

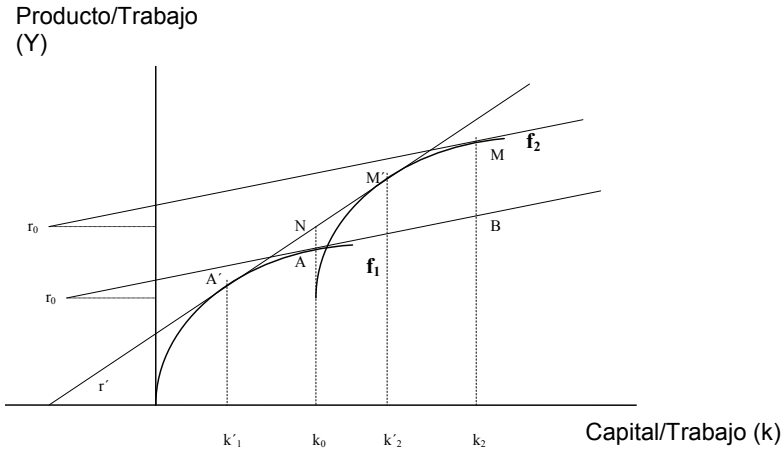
En realidad, lo más frecuente no es observar una tecnología única en cada momento del tiempo, sino que la realidad se caracteriza por tecnologías heterogéneas que se utilizan de manera contemporánea.

El desarrollo formal y las implicancias empíricas de la heterogeneidad tecnológica fueron presentados originalmente por Mundlak (1988, 1993). Este autor supone que existe un conjunto de tecnología disponible (el “estado del arte”) y el cambio tecnológico es el cambio en este conjunto. Esta tecnología disponible contiene un subconjunto de tecnologías que no son implementadas y por lo tanto no observadas directa o indirectamente, lo que implica que la medición empírica del cambio en el conjunto de tecnologías disponibles no es posible. Cualquier medición debe restringirse a las tecno-

logías implementadas, que también son un subconjunto de las tecnologías disponibles. La diferenciación entre las tecnologías disponibles e implementadas no es trivial si existen varias técnicas disponibles. En este caso la elección de las técnicas implementadas puede afectar el cálculo en la productividad total de los factores (PTF).

Para ilustrar este punto la Figura 1 presenta dos funciones de producción, una que podemos llamar tradicional (f_1) y la otra moderna o nueva en el conjunto de tecnologías disponibles (f_2).

Figura 1. Tecnología Heterogénea



El eje horizontal mide la relación de insumos, capital por unidad de trabajo, y el eje de ordenadas mide el producto por unidad de trabajo. Inicialmente solo la tecnología tradicional está disponible y el producto está en el punto A con una relación k_0 de insumos. La respuesta a la aparición de la nueva tecnología puede ser diversa dependiendo de las condiciones de mercado y de las restricciones sobre las ofertas de factores. Si el sector es tomador de precios, la producción cambia desde el punto A al punto M , donde la relación capital-trabajo es

k_2 . El cambio total en el producto, Y_M/Y_A puede descomponerse en el efecto insumo, Y_B/Y_A , y el cambio relativo en la PTF, Y_M/Y_B . El punto Y_B se obtiene extendiendo la línea tangente a la función de producción en el punto A hasta el punto B, donde la relación capital/trabajo es k_2 .

Por otra parte, si la oferta de capital en el momento inicial es perfectamente inelástica, el ratio capital/trabajo permanece en k_0 . Pero no necesariamente el producto debe continuar en el punto A. Los recursos pueden asignarse en alguna proporción a cada una de las tecnologías y podría producirse el nivel de producto N, dado por alguna combinación lineal de f_1 y f_2 . Es decir, coexisten las dos tecnologías y se produce una mejora en la PTF dada por Y_N/Y_A . En la medida en que más capital está disponible (o se acumula), el movimiento se dará entre N y M' , a lo largo de la línea tangente a los puntos A' y M' . El movimiento a partir de N se da exclusivamente por el incremento en el uso de insumos, con lo cual no hay cambios en la PTF. Esto implica que el cambio en la PTF será menor que en el caso de oferta perfectamente elástica de capital. En este sentido, la estimación empírica de la PTF es condicional a las características de la oferta de recursos. El origen de las diferencias entre las estimaciones de PTF tiene que ver con la coexistencia de tecnologías y el cambio en los precios de factores o en su productividad marginal. La aparición de una nueva técnica que es más capital intensiva y también más productiva incrementa la demanda de capital. Cuando la oferta de capital no es perfectamente elástica, su precio (o su tasa de retorno) debe incrementarse para internalizar el cambio tecnológico. Específicamente, cuando el capital es inicialmente fijo, el movimiento posterior desde N a M' es totalmente explicado por el incremento en el capital utilizado. Entonces en el primer caso la contribución del insumo se obtiene utilizando la misma productividad marginal en la nueva y la vieja tecno-

logía (dada por el ratio de precios que no cambia). En el segundo caso, cuando las dos tecnologías coexisten, la productividad marginal del factor relativamente escaso se incrementa y la del otro factor declina. El cambio de ponderaciones entre factores absorbe parte del cambio técnico y se lo asigna a la retribución factorial. Este caso es relevante, ya que el cambio tecnológico puede ser de importante magnitud, pero si existen restricciones de oferta de factores y la tecnología utilizada termina siendo heterogénea el mismo puede estar subestimado. Este es el caso cuando el cambio tecnológico se sesga hacia el factor relativamente escaso y es interesante no sólo para el capital físico sino también para el caso del capital humano (Mundlak, 2000). Las implicancias empíricas de esta visión del cambio tecnológico son importantes ya que permiten explicar algunos fenómenos frecuentes en la producción agropecuaria a nivel sectorial, tales como las disparidades de productivas intra-sectoriales, las amplias brechas de productividad observadas entre productores y la dispersión en las tasas de adopción de nuevas tecnologías.

8. La Producción en Condiciones de Incertidumbre

En la presentación anterior del modelo de producción, la elección de los niveles de insumos se realizó bajo un esquema de certidumbre. Es decir, las distintas alternativas de decisión son conocidas y en consecuencia el decisor elige entre alternativas fijas o, de manera equivalente, no se considera la existencia del riesgo.

La incorporación del riesgo como característica importante en la toma de decisiones productivas implica modificar el modelo conceptual utilizado y también parte de sus extensiones, tales como la aplicación del enfoque de dualidad.

La introducción del riesgo en la teoría de la producción y sus aplicaciones empíricas se han basado en su mayor parte en la idea de la función de producción estocástica. De mane-

ra simplificada, la función de producción estocástica es una función de producción con un término aleatorio adicional, por ejemplo:

$$(24) \quad y = f(x) = x_1^\alpha x_2^\beta + \varepsilon,$$

donde y es el nivel de producto, $f(x)$ es la función de producción, x_1 y x_2 son insumos y ε es un término de error que representa el riesgo. En este caso, el riesgo o dispersión en los resultados creado por el término de error no afecta las decisiones marginales de utilización de insumos.

En una representación un poco más compleja, puede extenderse el caso para considerar el caso de que el riesgo sea función del nivel de utilización de los insumos:

$$(25) \quad y = \tilde{f}(x) = x_1^\alpha x_2^\beta + \varepsilon(x_1, x_2),$$

donde $\tilde{f}(x)$ es la función de producción estocástica, donde ahora el riesgo es función de la cantidad de factores utilizados.

Este enfoque de función de producción estocástica fue propuesto originalmente por Just y Pope (1978). Esta especificación permite a los insumos incrementar o disminuir el riesgo de producción. No obstante, el ratio de utilización de insumos no cambia debido a la forma particular en la cual el riesgo entra en la función de producción.

Alternativamente, es posible proponer una función de producción que depende de un conjunto de insumos determinísticos y de un insumo estocástico:

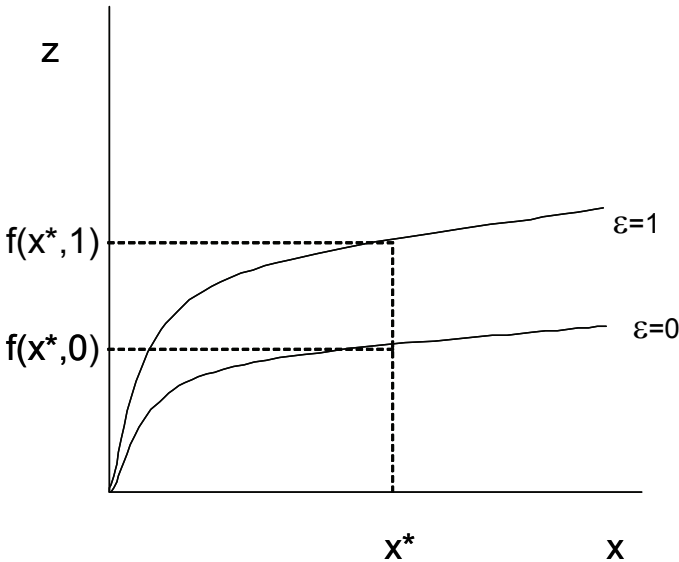
$$(26) \quad z = f(x, \varepsilon),$$

donde z es el nivel de producto, el cual es una función de los insumos que pueden ser controlados por la empresa, x , y factores no controlables o estocásticos ε .

Para simplificar la interpretación puede suponerse que el componente aleatorio es una variable Bernoulli que puede

tomar valores $\varepsilon = \{0,1\}$. Si este es el caso, la función de producción puede representarse como en la Figura 2, donde la variable aleatoria tiene un efecto que es independiente de la elección de insumos.

Figura 2. La función de producción estocástica.



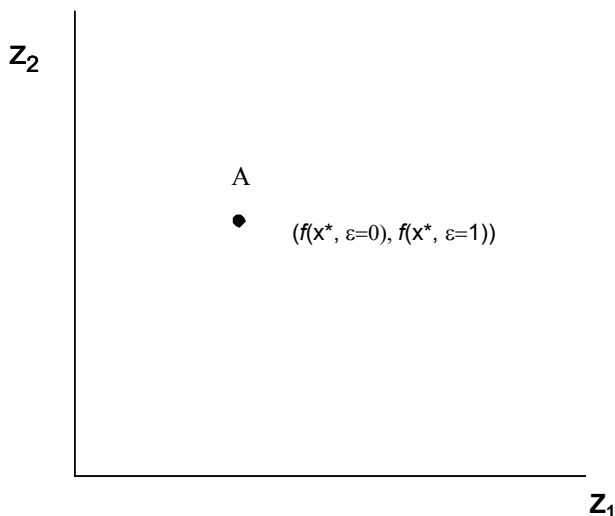
Es decir, la presencia del riesgo como variable aleatoria simplemente modifica la posición de la función de producción haciéndola condicional al valor que asume la variable ε .

9. El Enfoque de la Función de Producción Contingente

La especificación de la función de producción en las ecuaciones (24) y (25) es similar a la propuesta para el estudio del cambio tecnológico no incorporado. Este tipo de cambio tecnológico implica que no se introduce sesgo factorial o efectos sobre los insumos al producirse el cambio tecnológico. En lugar de esta aproximación, Chambers y Quig-

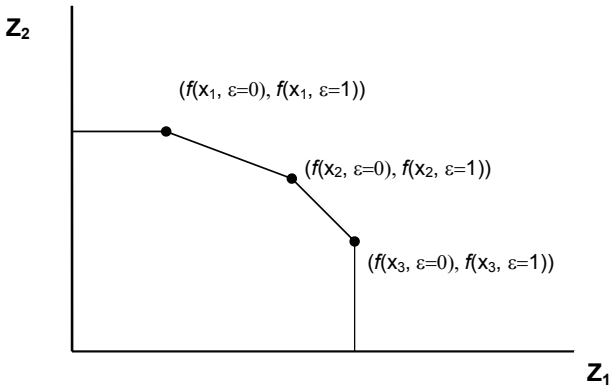
gin (2000) proponen describir la relación entre el uso de insumos y los eventos aleatorios en una dimensión diferente, el espacio de producción contingente al estado de la naturaleza (*state-contingent output space*). Estos autores utilizan como punto de partida las ideas desarrolladas por Arrow y Debreu para tratar el problema de la incertidumbre en el análisis de equilibrio general. La incertidumbre es representada por un conjunto de posibles estados de la naturaleza y los productos inciertos son vectores de bienes contingentes. Entonces, se supone que la producción de cada bien en cualquier momento es contingente al estado de la naturaleza que puede ocurrir. En este espacio, una decisión simple de producción implica dos diferentes estados posibles en términos de producción, tal como se representa por el punto A en el Figura 3.

Figura 3. Realizaciones de producción en dos estados de la naturaleza para un conjunto dado de insumos.



Entonces, diferentes elecciones de producción en términos de utilización de insumos generan diferentes puntos en el espacio tal como se ve en la Figura 4.

Figura 4. Posibles estados para diferentes niveles de uso de insumos



Las elecciones de producción pueden denominarse *acciones* (a) que tienen diferentes consecuencias (y) en diferentes estados (s). El conjunto de posibles estados se define como $\Omega = \{1, 2, \dots, S\}$, el conjunto de posibles consecuencias es $Y \subseteq \mathbb{R}^M$, y las acciones son una representación del efecto del conjunto de posibles estados sobre el conjunto de consecuencias $A: \Omega \rightarrow \mathbb{R}^M$.

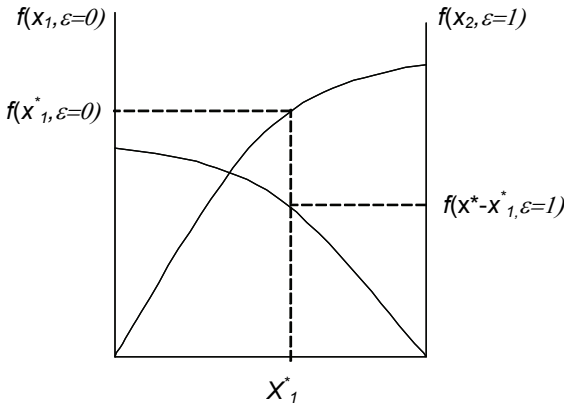
Cada acción define entonces un conjunto de vectores contingentes de resultados $y(A) \in \mathbb{R}^{M \times S}$. Debe notarse que se supone que el tomador de decisiones elige las acciones teniendo en cuenta las potenciales consecuencias para cada estado. En este enfoque el conjunto de estados de la natura-

leza es una descripción de los todos posibles estados, mutuamente exclusivos y exhaustivos.⁵

Es importante remarcar que lo que caracteriza esta aproximación al problema es que la cantidad de insumo aplicada es dependiente del estado (en términos de Chambers y Quiggin “*state allocable input technology*”). En el caso de la variable aleatoria definida antes (ε) supongamos, por ejemplo, que en una función de producción de un cultivo agrícola si la variable es $\varepsilon=0$, significa sequía mientras que $\varepsilon=1$ una cantidad de lluvia adecuada, aunque no necesariamente óptima. Es decir, en condiciones normales de lluvias durante el período de desarrollo del cultivo el agua puede administrarse realizando algunas acciones o prácticas agrícolas, las cuales permiten incrementar la producción. En períodos de sequía, estas acciones deberían ser reemplazadas por gastos en irrigación o en mitigación de la falta de agua, lo cual también permitirá incrementar la producción en este estado de la naturaleza. Supongamos que hay una cantidad total fija de esfuerzo o recursos (x^*) destinados a la mitigación de las condiciones climáticas, sea por exceso de agua o sequía. Ante un clima incierto, el total de esfuerzo disponible debe destinarse, ex ante, a las dos actividades posibles de mitigación, es decir $x_1^* + x_2^* = x^*$. Estos insumos contingentes pueden sustituirse uno por otro tal como se representa en la Figura 5.

Figura 5. Decisiones de uso de insumos contingentes al estado de la naturaleza

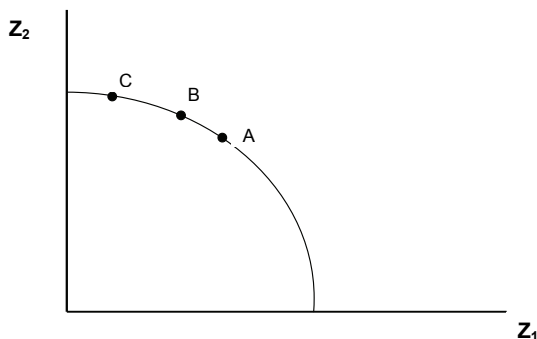
⁵ Esto desde el punto de vista práctico impone ciertas limitaciones para la aplicación empírica ya que el conjunto posible puede ser muy grande y diverso. Por ejemplo, la definición o formas de describir el clima y todas las variantes o permutaciones que pueden ocurrir en el mismo.



El eje horizontal mide el total de esfuerzo disponible. De izquierda a derecha se mide el esfuerzo destinado a actividades de mitigación en caso de sequía, mientras que el destinado a administrar el agua se mide de izquierda a derecha. El eje horizontal izquierdo mide el nivel de producción cuando $\varepsilon=0$, mientras que el eje derecho mide la producción si $\varepsilon=1$. Las funciones de producción, ex post, $f(x_1, \varepsilon=0)$ y $f(x_2, \varepsilon=2)$ se grafican respectivamente en los ejes izquierdo y derecho mostrando la productividad marginal decreciente del esfuerzo en cada estado de la naturaleza.

Supongamos que inicialmente el nivel de esfuerzo x_1^* se aplica a actividades de irrigación mientras que el complemento $(x^* - x_1^*)$ se aplica a acciones de administración del agua. Entonces, en el estado de la naturaleza $\varepsilon=0$ la producción es $f(x_1^*, \varepsilon=0)$ y en el caso de $\varepsilon=1$ el producto es $f(x^* - x_1^*, \varepsilon=1)$. Es decir, el insumo es contingente a la ocurrencia del estado de la naturaleza. De forma tal que en el caso de que ocurra el estado de la naturaleza i , dado que se destinan recursos para mejorar el producto en ese estado, es posible obtener funciones de producción ex post específicas para cada uno de los estados. Esto puede representarse en el espacio de producción contingente como el punto A del Figura 6.

Figura 6. Cambio en el uso de insumos en el espacio de producción contingente



Si, por ejemplo, se reasigna (ex ante) parte del esfuerzo de la actividad de irrigación (x_1) hacia la de administración del agua (x_2), entonces cuando ocurre $\varepsilon=0$ el producto disminuye, mientras que si $\varepsilon=1$ el producto aumenta. En la Figura 6 esto puede representarse como un movimiento de A hacia B. Si este experimento se repite en ambas direcciones, la reasignación del esfuerzo entre x_1 y x_2 permite derivar un conjunto de producción contingente que muestra la transición entre los diferentes estados, tal como se muestra en la Figura 6, y puede obtenerse una curva de transformación de estados contingentes cóncava al origen. La pendiente negativa de esta frontera puede pensarse como la tasa marginal de transformación entre estados contingentes, mientras que la concavidad refleja el hecho de que incrementar el producto en un estado contingente se realiza implica un costo creciente en términos de producción en el otro estado contingente.

A diferencia de este caso, el enfoque de la función de producción estocástica resultaría en un solo punto en el espacio, dado que disminuir el uso del insumo en uno de los estados de la naturaleza no permite incrementar el producto en otro estado, es decir, los insumos no son específicos al estado de la naturaleza.

Por ejemplo, bajo el enfoque de producción contingente en el caso de la agricultura mencionado antes, el insumo genérico “esfuerzo” aplicado a la irrigación tiene un efecto positivo sobre la producción en caso de sequía. Alternativamente, el esfuerzo puede dedicarse a trabajos de drenaje o mitigación en caso de exceso de lluvia, lo cual también tiene productividad marginal positiva. Sin embargo, el esfuerzo a irrigación tiene efectividad marginal cero ex post en caso de exceso de lluvia. Bajo el enfoque de función de producción estocástica el insumo no tiene nunca productividad marginal cero ex post.

De esta manera, es posible describir el conjunto posible de estados de la producción que pueden ser generados por un vector de insumos. Luego, también es posible derivar funciones de comportamiento económico tal como se realiza con las funciones de producción usuales a través del enfoque dual.

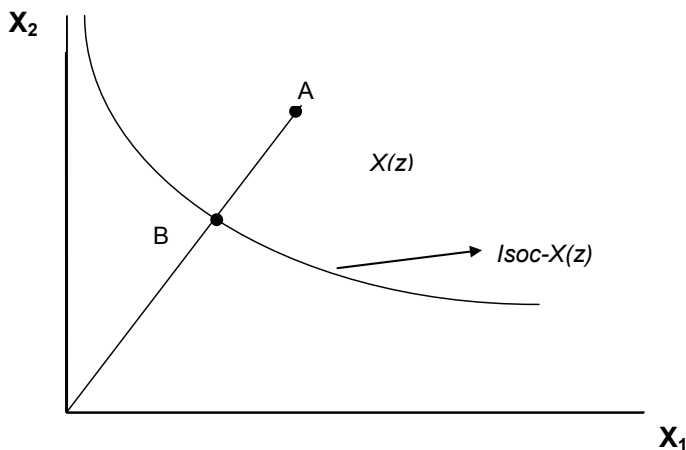
También es posible considerar bajo esta aproximación la existencia de ineficiencia técnica en la producción. Las posibilidades técnicas están representadas en todos los puntos sobre y por debajo de la frontera de la Figura 6. Cualquier punto por debajo de la frontera puede considerarse técnicamente ineficiente y pueden aplicarse las medidas usuales de estimación de eficiencia.

Por ejemplo, un instrumento utilizado en el análisis de producción teórico y aplicado al estudio de productividad son las funciones de distancia. En la teoría de la producción, una función de distancia orientada a los insumos se define como la fracción del insumo que puede ser retirada del proceso productivo de forma tal que aún se satisfaga un determinado nivel de producción. Formalmente:

$$(27) \quad D_I(x, z) = \max_{\theta} \left\{ \theta : \left(\frac{x}{\theta} \right) \in X(z) \right\},$$

donde θ es un escalar de distancia, x es un vector de insumos y $X(z)$, es un conjunto de insumos que puede producir al menos el producto z . El supuesto usual es que $\theta \geq 1$ es decir que la frontera es el mejor resultado posible en términos de eficiencia. Un ejemplo de función de distancia se representa en la Figura 7, donde x_1 y x_2 son insumos que se utilizan para producir el vector de productos z .

Figura 7. La función de distancia y el conjunto de requerimiento de insumos.



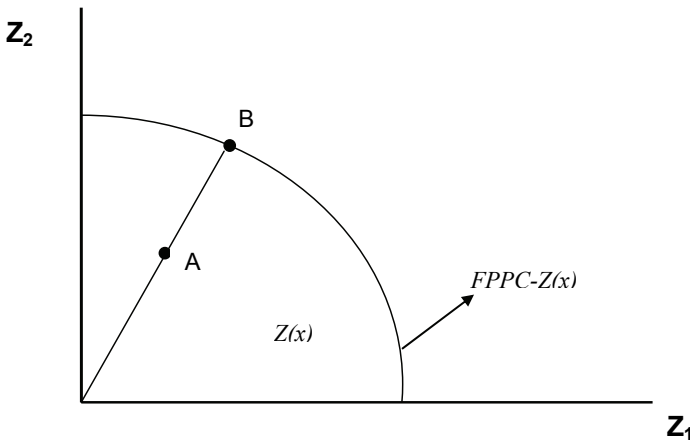
El conjunto de requerimientos de insumos, $X(z)$, es el área que está limitada por debajo por la isocuanta $Isoc-X(z)$. El valor de la función de distancia para el punto A, el cual define el punto de producción donde la empresa A utiliza las cantidades de insumos x_{1A} y x_{2A} para producir el producto z , se define como el cociente $OA/OB = \theta$.

De manera similar Chambers y Quiggin proponen una función de distancia contingente orientada al producto como:⁶

$$(28) \quad O(z, x) = \min\{\theta > 0 : \frac{z}{\theta} \in X(z)\}$$

La Figura 8 muestra una función de distancia en estados contingentes utilizando el ejemplo de dos productos, z_1 y z_2 , para un vector de insumos x .

Figura 8. La función de distancia y el conjunto de posibilidades de producción contingente.



Para un conjunto dado de insumos podemos representar la tecnología de producción mediante el conjunto $Z(x)$ que es el área delimitada por la frontera de posibilidades de producción contingente, $FPPC-Z(x)$ y por los ejes z_1 y z_2 . El valor de la función de distancia para la empresa que utiliza el nivel de insumos contingentes x para producir los productos definidos por el punto A es equivalente al cociente

⁶ La definición de la función de distancia podría ser más rigurosa reemplazando el “min” (*minimum*) con “inf” (*infimum*). Esto permitiría la posibilidad que el mínimo no exista (es decir que $\theta = \infty$ es posible).

OA/OB = $\tilde{\theta}$ Esta medida de distancia es menor o igual a uno y mide la inversa del factor por el cual todas las cantidades de productos podrían ser incrementadas, bajo la condición de permanecer dentro del conjunto factible de producción para el nivel de insumos determinado. Entonces, este tipo de medidas usualmente utilizadas para la estimación de eficiencia pueden ser también aplicadas dentro del enfoque de producción contingente para medir la eficiencia relativa de la toma de decisiones bajo incertidumbre.

Es posible también utilizar el enfoque de dualidad y pueden derivarse las funciones de beneficios y costos para realizar análisis de estática comparativa y aplicar el conjunto de técnicas disponibles en la teoría de la producción. Estas técnicas tienen aplicación directa en este enfoque, a diferencia de la función de producción estocástica, ya que la función de producción en estados contingentes puede presentarse como un caso particular del caso general de producción con múltiples insumos y productos (Chambers y Quiggin 2006).

10. Aplicaciones del Enfoque de Producción en Estados Contingentes

Chambers y Quiggin (2000) sugieren que esta aproximación provee la mejor manera de pensar sobre los problemas económicos bajo incertidumbre, incluyendo problemas de elección del consumidor, teoría de la firma y relaciones de principal-agente. Así, en los últimos años una importante literatura aplicada se ha desarrollado utilizando este enfoque para analizar problemas económicos donde deben modelarse aspectos productivos tales como seguros de cultivos (Chambers y Quiggin 2002), reformas de seguridad social (Grant y Quiggin 2002), contaminación (Chambers y Quiggin 1997), políticas de mitigación de riesgo de sequías (Quiggin y Chambers 2004) problemas de contaminación de fuentes pun-

tuales (Quiggin y Chambers 1998), biopropectiva (Smith y Kumar 2002) y regulación bancaria (Suwandi 1995).

Las aplicaciones empíricas no son tan frecuentes. En parte porque existen dificultades en estimar tecnologías de producción cuando las respuestas de los productores a la incertidumbre son endógenas. La aplicación de los métodos estándar de tecnologías de producción multi-producto al problema de incertidumbre requeriría de observaciones completas de los vectores (x, z) . Sin embargo, dado que cada observación se asocia con una realización de un particular estado s , los puntos observados son de la forma (x, zs) . Es decir, buena parte de los datos que serían necesarios para la aplicación de los métodos estándares no está disponible ya que se encuentra “perdida” en los estados de la naturaleza potenciales y no realizados.

Desde el punto de vista teórico, para una completa representación del problema, tanto el número de estados de la naturaleza como el número de momentos para describir la producción podrían ser infinitos (o al menos mayor que el número de grados de libertad provisto por un conjunto de datos factible).

Por lo tanto, en el trabajo empírico el número de estados de la naturaleza que pueden incluirse es relativamente pequeño, dos o tres, de manera similar a lo usual en los modelos estocásticos que trabajan con los momentos de primer y segundo orden (media y varianza) para describir el comportamiento de la producción.

Además de este problema de elección del número de estados de la naturaleza, también debe resolverse el hecho de que sólo uno de los estados se realiza para cada observación. Griffiths y O’Donnell (2004) utilizan un enfoque de máxima verosimilitud para resolver esto en el contexto de de un modelo de frontera de producción en el cual el producto puede estar por debajo del óptimo técnico como consecuencia de un efecto de ineficiencia específico para la firma. La tecno-

logía subyacente se supone contingente y las observaciones se asignan a uno de tres posibles estados de la naturaleza elegidos en base a un criterio de máxima verosimilitud.

Un resultado interesante de esta aproximación es que comparados con modelos estándar de fronteras estocásticas, los efectos de ineficiencia son mucho menores. Las observaciones, que serían ineficientes en un modelo de frontera tradicional, pueden ser en cambio representadas como el resultado de un estado de la naturaleza desfavorable.

Chambers y Quiggin (2004) utilizan también esta aproximación para definir indicadores estocásticos de productividad aplicados a datos de la agricultura de posguerra en los EE.UU. de Norteamérica. Las comparaciones con otros indicadores de productividad muestran que la consideración del entorno estocástico en el cual operan las empresas agropecuarias tiene importantes implicancias en el resultado cuantitativo de la medición de productividad.

En una aplicación empírica reciente, Moss (2010, capítulo 11) muestra que una formulación de la función de producción estocástica que es consistente con el modelo de producción en estados contingentes puede ser derivada de una regresión por cuantiles⁷ en la versión propuesta por Koenker y Bassett (1978). Este autor, propone la especificación de una función de producción cuadrática utilizando regresión por cuantiles y realiza la estimación para la producción de trigo en Kansas, Missouri, Nebraska, Oklahoma y Texas en función de la aplicación de los insumos nitrógeno y fósforo utilizados como fertilizantes. A partir de los cambios en los términos lineales y cuadráticos de la estimación este autor deriva modelos de producción contingentes que le permiten comprobar que la aplicación óptima de fertilizantes varía dependiendo del esta-

⁷ Para una revisión y síntesis de la metodología de la regresión por cuantiles puede verse Sosa Escudero 2005.

do de las variables no controladas, tal como sugiere la teoría de producción en estados contingentes.

11. Síntesis y Comentarios Finales

La literatura teórica y empírica sobre la producción, en particular en temas de economía agropecuaria, ha evolucionado en los últimos 60 años desde aspectos específicos vinculados con la función de producción en si misma hacia el análisis conjunto que trata de vincular las condiciones tecnológicas con la optimización económica y el comportamiento bajo condiciones de incertidumbre.

El enfoque inicial fue el de la estimación directa de funciones de producción de tipo Cobb-Douglas utilizando datos micro y macroeconómicos. Los aportes más importantes de este enfoque consistieron en la provisión de una metodología para estimar elasticidades, testear las condiciones de optimización derivadas de la maximización de beneficios y en la medición de las economías de escala. Sin embargo, los resultados en general han sido variables y condicionales al conjunto de datos utilizados. Además, se han señalado los sesgos estadísticos generados por la correlación entre factores idiosincrásicos no observables y los niveles de uso de insumos. Es decir, el problema de endogeneidad de los insumos en las estimaciones directas de funciones de producción.

Una aproximación general para solucionar este problema ha sido la utilización de los precios como variables instrumentales y el aprovechamiento de la variabilidad en datos de panel. Asimismo, la incorporación en el análisis teórico de formas funcionales más complejas y flexibles que la función Cobb-Douglas (por ejemplo funciones tipo CES, Cuadráticas o Translogarítmicas) ha implicado también una mayor dificultad de la estimación econométrica directa, en muchos casos debido al fuerte problema de multicolinealidad

como consecuencia de la mayor cantidad de parámetros a estimar.

Una forma de solucionar estos problemas ha sido la imposición de las condiciones competitivas y realizar luego la estimación a partir del uso de las condiciones de primer orden para la maximización de beneficios. Esta solución, si bien permite la estimación, implica la imposibilidad de testear las condiciones de competencia.

El interés central en la estimación directa de las funciones de producción está fundado, en general, en permitir el cálculo de las elasticidades de las funciones de comportamiento de interés económico, tales como la oferta de productos, la demanda de factores y el valor e las funciones objetivo de costos, beneficios o ingresos.

Los avances teóricos y la aplicación general del enfoque de dualidad permitieron un camino inverso para la estimación empírica, generando el punto de partida en la función objetivo. Bajo este enfoque, si la función objetivo es conocida, es posible recuperar los parámetros relevantes de la función de producción implícita. El uso de la dualidad tanto en la teoría como en la aplicación empírica resulta a partir de los años setenta muy atractivo por diversas razones: i) es un concepto teórico relevante y poderoso; ii) los precios se suponen exógenos y en consecuencia pueden ser utilizados en las estimaciones superando el problema de la endogeneidad y iii) permite la identificación de la tecnología subyacente. No obstante algunos problemas persisten aun en este enfoque ya que el punto ii) sólo sería válido en estimaciones con micro datos y el punto iii) sólo si la tecnología efectivamente implementada es independiente de los precios (Mundlak 2000).

Los avances en técnicas de estimación econométrica han popularizado este enfoque y también han permitido mejoras en las estimaciones para controlar el problema de endogeneidad de las variables explicativas. Sin embargo, la duali-

dad entre la tecnología y los precios se sostiene en formas funcionales flexibles que tienen solución cerrada bajo condiciones bien definidas que pueden ser puestas a prueba empíricamente.

En muchos de los estudios empíricos estas condiciones no suelen cumplirse. En particular la concavidad de la función de costos o la convexidad de la función de beneficios, lo que implica inconsistencias entre la tecnología estimada y los supuestos del modelo.

La heterogeneidad tecnológica observada en la producción real también es una dificultad que plantea desafíos para la estimación empírica. La conceptualización del problema de heterogeneidad tecnológica sugiere que la implementación de las tecnologías disponibles es gobernada por consideraciones económicas (tecnología endógena) y en consecuencia es afectada por las variables utilizadas en el análisis convencional tales como los precios de productos e insumos. Esta aproximación sugiere también una interesante vía de investigación para analizar los potenciales efectos sobre la productividad de las políticas públicas que afectan precios relativos de productos e insumos.

Finalmente, la introducción de incertidumbre implica nuevos problemas conceptuales para la representación de la tecnología en un contexto estocástico. En un contexto determinístico, la teoría de la producción tiene una equivalencia formal con los desarrollos en teoría del consumidor, lo que hace que los instrumentos de un área puedan ser aplicados muchas veces en la otra y también permiten una analogía que facilita la comprensión y la intuición. La teoría de la dualidad es en este sentido ilustrativa. Por ejemplo, la simetría entre la función de gasto del consumidor y la función de costo muestra cómo las herramientas utilizadas en un contexto pueden aplicarse a otro de manera relativamente simple.

Esta analogía no se mantiene cuando se incorpora la incertidumbre al análisis. Los modelos de consumo y producción bajo incertidumbre difieren considerablemente en su aproximación al caso determinístico y también entre ellos en contextos estocásticos.

El análisis de los problemas de producción bajo incertidumbre requiere un nuevo conjunto de técnicas y conceptos. De forma usual el problema se representa con una función de producción estocástica que responde de alguna manera a fuentes exógenas de incertidumbre. Este enfoque implica que ya no se dispone, por ejemplo, de las herramientas gráficas conocidas desde el modelo determinístico, tales como isocuantas, curvas de preferencias y curvas de transformación. El tratamiento de los problemas es altamente analítico y complejo. Un corolario derivado de esta aproximación es que la dualidad ya no es aplicable bajo incertidumbre.

Esto plantea también un problema epistemológico. La razón de estudiar la producción en un contexto determinístico es tratar aprender algo sobre el mundo real que por definición es estocástico. Si las herramientas desarrolladas bajo la abstracción no estocástica no son utilizables en un contexto incierto, entonces podría cuestionarse el valor de esta aproximación para comprender algunos problemas reales.

El enfoque de la producción en estados contingentes desarrollado recientemente por Chambers y Quiggin (2000) es una forma alternativa de tratar el tema que permite utilizar las herramientas desarrolladas en el enfoque determinístico de manera coherente en un contexto estocástico. La idea es utilizar el concepto desarrollado por Arrow y Debreu de bienes en estados contingentes para la teoría del equilibrio general. Este concepto diferencia los bienes de acuerdo al momento y lugar en el cual son entregados. Es decir, los bienes son contingentes a la ocurrencia de un determinado

estado de la naturaleza. Esta forma de analizar el problema permite obtener una teoría de la producción bajo incertidumbre libre de complejidad probabilística y que permite utilizar el instrumental básico de la teoría bajo certidumbre. Así, en los últimos años se han desarrollado diversas aplicaciones empíricas del enfoque de producción contingente, en particular luego del trabajo de Griffiths y O'Donnell (2004) que estimaron un modelo de frontera de producción con una tecnología subyacente que se supone contingente. Este nuevo enfoque teórico, y sus aplicaciones empíricas, son un progreso reciente que según Quiggin y Chambers (2006) puede ser aplicado de manera conveniente en casi cualquier problema económico que implique incertidumbre.

12. Referencias bibliográficas

- Antle, J. M. (1984) "The structure of US agricultural technology, 1910-1978" *American Journal of Agricultural Economics*, 66(4):414-421.
- Chambers, R.G. and J. Quiggin (2000) "Uncertainty, Production, Choice and Agency: The State-Contingent Approach" Cambridge University Press, New York.
- Chambers, Robert G. (1988). "Applied production analysis: a dual approach" Cambridge University Press.
- Chambers, R.G. and Quiggin, J. (1997) "Separation and hedging results with state-contingent production", *Economica* 64, 187-209.
- Chambers, R.G. and Quiggin, J. (2002) "Optimal producer behavior in the presence of areayield crop insurance" *American Journal of Agricultural Economics* 84, 320-334.
- Chambers, R.G. and Quiggin, J. (2004) "Technological and financial approaches to risk management in agriculture:

- an integrated approach” *Australian Journal of Agricultural and Resource Economics* 48, 199–223.
- Christensen, L.R., D. Jorgenson, and L. Lau (1971) “Conjugate Duality and the Transcendental Logarithmic Production Function”, *Econometrica*, 39(4): 255-256.
- Diewert, W. E. (1971), “An Application of the Shepard Duality Theorem: A Generalized Linear Production Function”, *Journal of Political Economy*, 79(3): 482-507.
- Diewert, W. E. (1974), “Applications of Duality Theory”, en M. Intriligator and D. A. Kendrick, eds., *Frontiers of Quantitative Economics*, Vol. II, North Holland, Amsterdam.
- Fulginiti, L. and R. Perrin (1990) “Argentine Agricultural Policy in a Multiple-Input, Multiple-Output Framework” *American Journal of Agricultural Economics*, 72(2): 279-288.
- Fulginiti, L. and R. Perrin (1993) “The theory and measurement of producer response under quotas” *The Review of Economics and Statistics*, 75(1).
- Grant, S. and Quiggin, J. (2002). “The risk premium for equity: implications for the proposed diversification of the social security fund” *American Economic Review* 92, 1104–1115.
- Griffiths, W. and O’Donnell, C. (2004) “Estimating State-Contingent Production Frontiers” Department of Economics, Research Paper Number 911, University of Melbourne, Melbourne.
- Heady, E. O. and J.L. Dillon (1961) “Agricultural Production Functions”. Iowa State University Press, Ames.
- Heady, Earl O. (1946) “Production functions from a random sample of farms”, *Journal of Farm Economics* 28(4): 989-1004.

- Jorgenson, D. W. (1986) “Econometric methods for modeling producer behavior” en Z. Griliches and M.D. Intriligator, eds. *Handbook of Econometrics*, Vol. III, North Holland, Amsterdam, 1841-1915.
- Just, R. E. and R. D. Pope (1978) “Stochastic specification of production functions and economic implications” *Journal of Econometrics* Vol.7(1): 67-86.
- Klein, L. R. (1953), “A Textbook of Econometrics”, Row, Peterson and Co., Evanston IL.
- López, R. E. (1984) “Estimating substitution and expansion effects using a profit function framework” *American Journal of Agricultural Economics*, 66(3):358-367.
- Mc. Fadden (1978), “Cost, revenue and profit functions”, en M. Fuss and D. McFadden eds. *Production Economics: A Dual Approach to Theory and Applications*, North Holland, Amsterdam, 3-109.
- Moss, Charles B. (2010) “Risk, uncertainty and the agricultural firm” World Scientific Publishing Co., Singapore.
- Mundlak, Yair (1993) “On the empirical aspects of economic growth theory”, *American Economic Review*, 83(2):415-420.
- Mundlak, Yair. (1988). “Endogenous technology and the measurement of productivity” en S. M. Capalbo and J.M. Antle, eds *Agricultural Productivity: Measurement and Explanation. Resources for the Future*, Washington DC, 316-33.
- Mundlak, Yair. (2000). “*Agriculture and Economic Growth*” Harvard University Press. Cambridge, Massachusetts.
- Quiggin, J. and R. G. Chambers (2006) “The state-contingent approach to production under uncertainty”

The Australian Journal of Agricultural and Resource Economics, 50:153–169.

Quiggin, J. and Chambers, R.G. (1998) “Risk premiums and benefit measures for generalized expected utility theories” *Journal of Risk and Uncertainty* 17, 121–137.

Quiggin, J. and Chambers, R.G. (2004) “Drought policy: a graphical analysis” *Australian Journal of Agricultural and Resource Economics* 48, 225–251.

Shumway, C. R. (1995) “Recent duality contributions in production economics”, *Journal of Agricultural and Resource Economics*, 20(1):178-194.

Smith, R.B. and Kumar, P. (2002). Royalties and Benefit Sharing Contracts in Bioprospecting. Working Paper E/221/2002, Institute of Economic Growth, Environmental Economics Unit, University Enclave, Delhi.

Sosa Escudero, W. (2005) “Perspectivas y avances recientes en regresión por cuantiles” en Mariana Marchioni (editora), *Progresos en Econometría*, Cap. 5. Serie Progresos en Economía, AAEP.

Suwandi, T. (1995). Indonesian Banking Post-deregulation: Moral Hazard and High Real Interest Rates. Australian National University, Canberra.

Tintner, G (1944), “A note on the derivation of production functions from farm records” *Econometrica* 12:26-34

Tintner, G. and O. H. Brownlee (1944), “Production functions derived from farm records” *Journal of Farm Economics* 26(3):566-571.

Economía de las Organizaciones

Christian Ruzzier
(Universidad de San Andrés)

1. Mercado versus organización

Al menos desde los tiempos de Adam Smith, con su descripción de la fábrica de alfileres, somos (¡los economistas, aunque más no sea!) plenamente conscientes de las ventajas de la especialización y la división del trabajo. En un mundo especializado en el cual nadie produce todo lo que desea consumir, los individuos se vuelven dependientes los unos de los otros. Aparece entonces la necesidad de coordinar las actividades de todos ellos y también de motivarlos, ya que no es inmediatamente obvio que esté en el interés individual el promover una solución eficiente al problema de coordinación (Roberts, 2004).

Comenzando con la publicación de *La Riqueza de las Naciones*, la Economía ha permanentemente destacado las virtudes del sistema de precios en esta tarea de coordinar y motivar la actividad humana: bajo determinadas condiciones, podemos dejar a cada quien perseguir sus propios fines egoístas y todos serán guiados, como por una mano invisible, a promover el interés general (Arrow y Hahn, 1971). Cuando los mercados funcionan correctamente, los problemas de coordinación y motivación se resuelven de manera descentralizada, mediante el sistema de precios. La cita que sigue resume esta visión.

“The normal economic system works itself. For its current operation it is under no central control, it needs no central survey. Over the whole range of human activity and human need, supply is adjusted to demand, and production to consumption by a process that is automatic, elastic and responsive.”

Sir Arthur Salter (citado en Coase, 1937)

Los economistas, de hecho, nos hemos vuelto expertos en el estudio de los mercados. Paradójicamente, buena parte de la actividad económica ocurre dentro de organizaciones como la empresa –es decir, *fuera* del mercado (véase Simon, 1991, para una elocuente argumentación de este punto). Considérense los siguientes datos:

Se estima que menos de un tercio de las transacciones en EEUU se realiza a través del mercado.¹

En 2007, Wal-Mart tenía 1,8 millones de empleados.²

Cerca de un tercio del comercio internacional es intra-firma.³

En febrero de 2007 habían 41 empresas con un valor de mercado de más de 100 mil millones de dólares.⁴

Al decir de D.H. Robertson, encontramos “islas de poder consciente” (las empresas) en este “océano de cooperación inconsciente” (el mercado).⁵ En estas “islas”, no es el mercado el que asigna los recursos a través del sistema de precios, sino que los mismos son frecuentemente asignados de

¹ El dato es de McMillan (2002), citado en Roberts (2004).

² Ver Hart (2008).

³ Ver, por ejemplo, Clausing (2003) y Freund y Weinhold (2002) para datos de Estados Unidos.

⁴ Ver nuevamente Hart (2008).

⁵ “[We find] islands of conscious power in this ocean of unconscious cooperation like lumps of butter coagulating in a pail of buttermilk”. La cita aparece en Coase (1937).

acuerdo a las órdenes de unas pocas personas –los managers– lo cual hace necesario persuadir a los demás miembros de la organización para que acaten estas órdenes. A veces esto se logra formalmente, por ejemplo, mediante un esquema de participación en las ganancias. Otras veces, la gente necesita ser inspirada, persuadida, convencida, forzada, y hasta despedida, para lograr que las órdenes se cumplan. Muchas veces aparecen conflictos entre managers, o entre ellos y otras partes con un interés en la organización (accionistas, trabajadores, clientes). Si bien puede haber alguna especie de negociación *à la* Coase (1960), rara vez estas transacciones se resuelven mediante precios. Más bien, en estos casos el sistema de precios, en mayor o menor medida, parece ser simplemente dejado de lado, como expresara vívidamente el mismo Coase en otro de sus trabajos clásicos:

“For instance, in economic theory we find that the allocation of factors of production between different uses is determined by the price mechanism. The price of factor A becomes higher in X than in Y. As a result, A moves from Y to X until the difference between the prices in X and Y, except in so far as it compensates for other differential advantages, disappears. Yet in the real world, we find that there are many areas where this does not apply. If a workman moves from department Y to department X, he does not go because of a change in relative prices, but because he is ordered to do so.”

Ronald Coase (1937)

La típica imagen de la empresa en la mayoría de los libros de texto –una caja negra que simplemente genera una cierta cantidad de producto a partir de una cierta cantidad de insumo– podría inducir a creer que los economistas no se toman muy en serio lo que ocurre en una gran porción de la economía. Más allá de una larga tradición en contrario (como atestigua la cita de Pigou que sigue), los economistas *han* estado tomándose seriamente a las organizaciones, es-

pecialmente en años recientes, y han hecho un progreso considerable en la caracterización de las mismas.

“...it is not the business of economists to teach woolen manufacturers to make and sell wool, or brewers how to make and sell beer, or any other business men how to do their job. If that was what we were out for, we should, I imagine, immediately quit our desks and get somebody—doubtless at a heavy premium, for we should be thoroughly inefficient—to take us into his woolen mill or his brewery...”

Arthur C. Pigou (1922)

Este capítulo pretende introducir al lector a algunas de las ideas nuevas que el campo de la Economía de las Organizaciones ha estado produciendo, y que han empezado a tener impacto sobre otras partes de la disciplina, incluyendo comercio internacional, organización industrial, economía laboral y macroeconomía.⁶ La Economía de las Organizaciones se superpone con muchos otros campos de la economía. Así, por ejemplo, hay partes del capítulo que son complementos naturales de un típico curso de organización industrial o de economía laboral, o que pueden verse como una aplicación del material de un curso de teoría de los contratos. Si bien en la mayor parte del capítulo haremos referencia a la empresa —o la corporación, un tipo particular de empresa— buena parte de las ideas se aplica igualmente a otros tipos de organización, como las ONG, las *joint-ventures* y hasta el Estado (véase, por ejemplo, Simon, 1991).

La pregunta del millón

Como documentáramos en la sección precedente, la mayoría de las transacciones económicas parecen ocurrir fuera

⁶ Para un tratamiento mucho más comprehensivo del campo, ver el *Handbook of Organizational Economics*, de próxima edición (Gibbons y Roberts, 2011).

de los mercados, a pesar de las mentadas virtudes de estos últimos en las tareas de coordinación y motivación de la actividad humana. La pregunta, entonces, se autoimpone: ¿por qué? Podemos precisar esta “pregunta del millón” de la siguiente manera: si los intercambios voluntarios entre diferentes agentes funcionan tan bien, ¿por qué recurrimos tanto a las empresas para coordinar y motivar la actividad económica, en vez de usar el sistema de precios o la negociación entre individuos? El primero en plantearse seriamente esta pregunta fue Ronald Coase. En sus propias palabras:

“In view of the fact that while economists treat the price mechanism as a coordinating instrument, they also admit the coordinating function of the ‘entrepreneur’, it is surely important to enquire why coordination is the work of the price mechanism in one case and of the entrepreneur in another.”

Ronald Coase (1937)

La frase sintetiza la nueva perspectiva aportada por Coase (que le valiera la mitad de su premio Nobel en 1991): mercado y empresa no son más que dos mecanismos alternativos para gobernar una transacción. Esto puede parecer obvio, pero en la época sólo lo fue para Coase.⁷ El hecho de que su trabajo pasara desapercibido por casi 40 años, cuando fue redescubierto y extendido por Oliver Williamson (quien recibió el premio Nobel en 2009 por esta contribución), es evidencia fiel de lo adelantado a su tiempo que estaba el pensamiento del joven Coase de 1937.

Una respuesta a la pregunta del millón la dio Arrow:

“Organizations are a means of achieving the benefits of collective action in situations where the price system fails. [...]”

⁷ Es interesante el hecho de que el mismo Coase vea su propia contribución a la Economía como precisamente eso: puntualizar obviedades (como él mismo expresara al recibir el premio Nobel).

All of them...[share] the need for collective action and the allocation of resources through nonmarket methods.”

Kenneth Arrow (1974)

La respuesta de Arrow (cuya esencia puede rastrearse hasta el trabajo original de Coase) es simple: a veces, los mercados no funcionan correctamente. Por lo tanto, no pueden brindar una solución eficiente al problema de coordinación y motivación. En estas circunstancias, otros mecanismos podrían ser mejores – una vez más, Coase nos recuerda que la empresa es la principal alternativa (véase Roberts, 2004). Consecuentemente, buena parte del trabajo de los economistas sobre las organizaciones se ha enfocado tradicionalmente en la elección entre mercado y empresa de acuerdo a criterios de eficiencia; es decir, la pregunta clásica de Coase sobre los límites de la empresa acaparó durante mucho tiempo la atención del campo (véase Gibbons, 2005a, para una introducción a las teorías más relevantes).

Pero si las empresas surgen sólo cuando los mercados son suficientemente malos en coordinar y motivar, aquéllas típicamente serán *second-best*; es decir, darán la mejor solución factible al problema de coordinación y motivación, pero esta solución será peor que la mejor solución imaginable (Gibbons, 1999). El razonamiento de Coase y Arrow parece tener, entonces, cosas para decirnos no sólo sobre los límites de la empresa, sino también sobre su funcionamiento interno: ¿por qué las causas que hacen fallar al mercado no habrían de afectar también a las empresas? La implicancia lógica es, por lo tanto, que es improbable que las empresas sean “relojitos suizos” –es decir, *first-best* (Gibbons, 2003)– y que deben ellas mismas estar plagadas de ineficiencias. De hecho, en todos los modelos presentados en este capítulo la ineficiencia es el resultado típico de las acciones de miembros de la organización que buscan alcanzar racionalmente

sus propios objetivos. En este sentido, los modelos de la organización interna de las empresas que introducimos predicen que éstas serán complicadas, pero no misteriosas (“a mess but not a mystery”, Gibbons, 1999, pág. 146).

En este capítulo, antes que revisar la literatura tradicional sobre la teoría de la firma (el problema de sus límites), optamos, por el contrario, por revisar las contribuciones de la teoría económica a la estructura interna de la empresa, dados los límites de la misma. En atención a las restricciones de espacio, nos limitamos a un enfoque particular –el de “precios” (*pricing*)– que pasamos a explicar.⁸

La provisión de incentivos en las organizaciones

Si el sistema de precios falla, la inclinación natural del economista es querer “arreglarlo”, a la manera en que un impuesto pigouviano soluciona la externalidad de la contaminación (Pigou, 1932). La primera tarea consiste, entonces, en fijar los “precios” de las distintas acciones dentro de la organización, *à la* Pigou (Gibbons, 2010a) – de allí el nombre del enfoque.⁹ En esta tarea, podemos valernos de diversos instrumentos directos e indirectos, formales o informales, que iremos desgranando a medida que avance el capítulo. Claramente, los precios que fijemos deben resultar menos que perfectos (en el sentido de inducir acciones de *first best*) –de no ser así, podríamos simplemente limitarnos a intervenir selectivamente en aquellos casos en que el mercado falla, y la implicancia lógica sería que deberíamos observar una única gran empresa, englobando toda la actividad económica (Williamson, 1985).

⁸ Véase Gibbons (2010a) para éste y otros enfoques en la teoría económica de la organización interna de la empresa.

⁹ No deja de haber cierta ironía en el hecho de que terminemos recorriendo a Pigou para estudiar temas que el propio Pigou nos recomendaba no encarar.

El modelo clásico de principal-agente con información asimétrica sobre las acciones (*moral hazard*) nos da el punto de partida. Si el principal y el agente comparten los objetivos, sin embargo, la no observabilidad de las acciones del agente no plantea ningún problema. El modelo básico plantea entonces cierto conflicto de intereses, necesario para tener incentivos imperfectos. Por sí solo, esto podría no ser suficiente: si el agente es neutral al riesgo y tiene riqueza ilimitada, su compensación puede estructurarse de tal manera que tome la decisión que tomaría el principal en su lugar. Ello requiere introducir riesgo en la compensación del agente, pero este riesgo no es costoso dado el supuesto de neutralidad. Si el agente fuera averso al riesgo, por otro lado, aparecería una disyuntiva entre el deseo de proveer incentivos fuertes y la necesidad de asegurar al agente; los incentivos resultan típicamente menos que perfectos en este caso, debido a la aversión al riesgo del agente. Repasamos brevemente este modelo en la sección 2.

Modelos más recientes han enfatizado una fuente alternativa de incentivos imperfectos: el uso de medidas distorsivas del desempeño. Mediante un contrato formal, el principal puede cometer “la insensatez de pagar por A esperando que se haga B” (“the folly of rewarding A while hoping for B”; Kerr, 1975) cuando contratar sobre la base de B se hace difícil. Un contrato basado en A puede generar incentivos muy poderosos... a realizar A, no necesariamente B. Aun si el agente es neutral al riesgo, aparece una potencial imperfección en los incentivos. En la sección 3 mostramos cómo esta distorsión afecta la provisión óptima de incentivos; es decir, analizamos cuándo tiene o no sentido pagar por A esperando B. También discutimos brevemente cómo el uso de otros instrumentos formales (como la propiedad de los activos y la limitación de las actividades que puede llevar adelante el agente) podría lograr mejores resultados.

La sección 4 resalta que, en muchas circunstancias, lo que puede lograrse con instrumentos formales es limitado, y llama la atención sobre el valor de los instrumentos informales. En particular, cuando las partes interactúan reiteradamente pueden recurrir a contratos implícitos (o relacionales): simples promesas cuyo valor radica en que pueden ser explicitadas en términos que un juez no entendería, pero que por el hecho de ser promesas deben ser formuladas de manera que esté en el interés de las partes cumplirlas. La sección 5, finalmente, reúne todos los instrumentos para discutir la idea de “sistema de incentivos”. Para ello, introducimos la noción de complementariedad entre instrumentos: dos instrumentos son complementarios cuando utilizar uno incrementa el rendimiento de emplear el otro. En presencia de complementariedades, el todo es más que la suma de las partes. La existencia de relaciones entre los distintos instrumentos tiene como consecuencia la aparición de un número restringido de patrones coherentes. El problema del diseño del sistema de incentivos consiste entonces en identificar y escoger entre varios patrones coherentes, ya que ningún instrumento aislado (formal, informal, implícito o indirecto) es capaz de crear buenos incentivos (Roberts, 2004).

2. El modelo simple de principal-agente con riesgo moral

Los esfuerzos individuales (físicos e intelectuales) suelen resultar en beneficios para toda la organización (en la forma de menores costos, mayores ventas, mejor reputación, etc.), pero frecuentemente son difíciles de observar y cuantificar en tiempo y forma. Como la organización no puede establecer a ciencia cierta lo que el individuo hizo o dejó de hacer, no es posible redactar un contrato que especifique lo que debe hacerse, así como los premios por cumplir el contrato y las penalidades por violarlo (Roberts, 2004). El modelo básico de principal-

agente (Holmström, 1979; Grossman y Hart, 1983) captura estas situaciones suponiendo la existencia de un agente que realiza una acción a que contribuye y al valor de la empresa (tradicionalmente la literatura llama *output* del agente a esta contribución), pero que no es observada por el principal. En este modelo, el *output* es de propiedad del principal, pero éste puede compartirlo con el agente mediante un contrato que establezca una compensación que dependa del *output*.

Siguiendo a Gibbons (2005b y 2010b), la cronología de eventos en el modelo básico sería:

1. El principal y el agente firman un contrato que especifica la compensación del agente como una función del *output*. Nos limitaremos a considerar contratos en los que la compensación total es una función lineal del *output*:

$$w(y) = s + b \cdot y,$$

donde s es el componente fijo de la compensación (por ejemplo, el salario) y b , el pago a destajo (es decir, por unidad de *output*).¹⁰

2. El agente elige una acción a , que el principal no observa. La interpretación más común de esta acción es que se trata del esfuerzo del agente, pero en general puede ser vista como cualquier acción que beneficia al principal y que el agente no realizaría en ausencia de una compensación. La razón por la que el agente demanda esta compensación es que la

¹⁰ Los contratos lineales son comunes en la práctica. Holmström y Milgrom (1987) proveen una justificación teórica de la optimalidad de los mismos. Bose et al. (2011), además, encuentran que generalmente el principal puede garantizarse al menos el 95% del beneficio que obtendría con un contrato óptimo mediante el uso de un contrato lineal.

acción le genera costos $c(a)$.¹¹ Por simplicidad, supondremos que $c(a) = a^2/2$ y que $c(0) = 0$.

3. Tienen lugar ciertos eventos fuera del control del agente (ε) y que afectan el output. Ello implica que hay ruido en la función de producción, lo que hace que el output del agente sea incierto.
4. La acción escogida por el agente y el ruido determinan el output. Supondremos que la tecnología de producción es $y = a + \varepsilon$ y que el ruido ε es una variable aleatoria que sigue una distribución normal con media cero y varianza σ^2 (un valor más alto de la varianza indica que hay más riesgo en el proceso productivo).¹²
5. El agente recibe la compensación especificada en el contrato, por lo que su compensación neta es $w(y) - c(a)$. El principal se queda con el output, pero debe pagar la compensación del agente, con lo que su pago neto es $y - w(y)$. Para comenzar, hacemos el supuesto de que tanto el principal como el agente son neutrales al riesgo, por lo que simplemente buscan maximizar el valor esperado de estos pagos.

Conviene comenzar por establecer cuál sería la acción de *first best*. La misma surge de maximizar el excedente total, $E[y - c(a)] = E[a + \varepsilon] - a^2/2 = a - a^2/2$. Claramente $a^{FB} = 1$. La siguiente pregunta es si el principal puede diseñar un contrato que induzca al agente a escoger esta acción; es decir, si el principal puede diseñar incentivos perfectos.

El problema es fácil de resolver por inducción hacia atrás. En $t = 2$ el agente resuelve el siguiente problema:

¹¹ Éste es el conflicto de intereses básico entre el principal y el agente.

¹² Por lo tanto, el output – al agregar el esfuerzo del agente y la pura suerte – es una señal imperfecta de la acción escogida por el agente.

$$\max_a E[w(y) - c(a)] = s + b \cdot a - \frac{1}{2}a^2,$$

cuya solución $a(b)$ nos da la respuesta óptima del agente frente al contrato ofrecido por el principal. De la condición de primer orden del problema del agente obtenemos:

$$(1) \quad a(b) = b$$

Nótese que el componente fijo de la compensación indicado como s no juega ningún rol en la provisión de incentivos, sino que es simplemente una transferencia de riqueza entre el principal y el agente. Uno de los atractivos de trabajar con un contrato lineal es que podemos resumir el poder de los incentivos en una única variable (b): un mayor b provee incentivos más poderosos, en el sentido de que induce más esfuerzo (o un nivel más alto de la acción) por parte del agente.¹³

En $t = 1$ el principal propone el contrato al agente. Dada la respuesta prevista del agente, el problema del principal se reduce a seleccionar el contrato que maximiza $E[y - w]$ sujeto a que el agente acepte dicho contrato. Ya que el principal tiene todo el poder de negociación en este momento, puede llevar al agente a su nivel de utilidad de reserva, que normalizamos en 0. Puede mostrarse entonces que este problema es equivalente a escoger b para maximizar el excedente total, $E[y(a(b)) - c(a(b))]$, o:

$$\max_b a(b) - \frac{1}{2} [a(b)]^2 = b - \frac{1}{2}b^2.$$

Claramente $b^* = 1$, lo que implica que $a^* = a^{FB} = 1$: con neutralidad al riesgo, el contrato óptimo induce acciones de *first best*. Mediante el contrato óptimo el principal efectiva-

¹³ La evidencia a este respecto es abrumadora; ver, por ejemplo, Lazear (2000).

mente “vende” la empresa al agente (transfiriéndole la propiedad sobre y) a cambio de un pago $s^* = (a^*)^2/2 - b^* \cdot a^* = -1/2 < 0$ con el que logra apropiarse de todas las ganancias ex ante.

$b^* = 1$ implica que el agente enfrenta todo el riesgo, mientras que $s^* < 0$ supone que tiene riqueza ilimitada (o al menos suficiente riqueza como para pagar s^* al comienzo de la relación). Esto sugiere dos vías para introducir imperfecciones en este modelo. La primera y más tradicional es suponer que el agente tiene aversión al riesgo; por lo tanto, debe ser compensado en este caso por el riesgo a que lo enfrenta el contrato. La segunda alternativa es suponer que el agente tiene restricciones financieras; es decir que no puede comprarle la empresa al principal.¹⁴ El principal elige en los dos casos reducir el poder de los incentivos (para reducir la prima de riesgo o la renta que se lleva el agente) y $b^* < 1$. Como las conclusiones cualitativas de ambas versiones son similares y ya que la disyuntiva entre riesgo e incentivos ha concentrado buena parte de la atención en la teoría de agencia, bosquejamos rápidamente el modelo con un agente averso al riesgo. Para ello, realizamos la modificación siguiente al modelo básico:

- 5'. El agente recibe la compensación especificada en el contrato, por lo que su compensación neta es $x = w(y) - c(a)$. El principal se queda con el output, pero debe pagar la compensación del agente, con lo que su pago neto es $y - w(y)$. El principal es neutral

¹⁴ Hay muchas formas alternativas de implementar el *first best*. La esencia de las mismas es recompensar los resultados buenos y penalizar los malos. Si el agente tiene riqueza limitada, hay una cota inferior a las penalizaciones. En ese caso, el principal sólo pueda usar las recompensas para inducir esfuerzo y debe darle cierta renta ex ante al agente (Lafont y Martimort, 2002).

al riesgo, por lo que simplemente busca maximizar el valor esperado de su pago. El agente, por su parte, tiene una función de utilidad $U(x) = -e^{-rx}$, en la que $r > 0$ es el coeficiente de aversión al riesgo absoluta del agente (mayor r implica que la aversión al riesgo es mayor).

El problema del agente es el de maximizar su utilidad esperada, $E[U(w(y) - c(a))]$, y la solución sigue estando dada por (1), pero el excedente total es ahora $a(b) - \frac{1}{2}[a(b)]^2 - \frac{1}{2}rb^2\sigma^2$, donde el último término representa la prima de riesgo que debe pagarse al agente para que esté dispuesto a aceptar el riesgo.¹⁵ Es claro que el principal querrá reducir b para minimizar esta prima. La solución de este problema es:

$$(2) \quad b^* = \frac{1}{1 + r\sigma^2}.$$

Claramente, $0 \leq b^* \leq 1$. Con $r > 0$, será óptimo fijar $b^* < 1$ – a menos que no haya ruido en la función de producción, en cuyo caso la no observabilidad del esfuerzo se vuelve irrelevante y el problema de motivación desaparece. El análisis de (2) revela que los incentivos serán débiles cuando el agente tenga una gran aversión al riesgo (r grande) o cuando haya mucha incertidumbre acerca del resultado de sus acciones (σ^2 grande). La disyuntiva entre brindarle incentivos al agente y forzarlo a enfrentar el riesgo lleva entonces a precios (e incentivos) menos que perfectos.

Si bien el modelo simple de principal-agente contiene importantes lecciones, también presenta serias limitaciones para comprender muchos fenómenos que se observan en el accionar cotidiano de las organizaciones. Por una parte, el

¹⁵ Consultar Gibbons (2005b) para los detalles.

modelo parece dejar fuera de toda consideración los problemas enfrentados por muchas empresas al tratar de introducir el pago por desempeño (ver los variados ejemplos de Kerr, 1975, y Baker et al., 1994). Además, sus resultados sugieren que si $b = 0$ no habría esfuerzo de parte del agente y, sin embargo, parece haber muchas situaciones en las que altos niveles de esfuerzo coexisten con algo como $b = 0$ (es decir, una situación en la que no hay un vínculo directo entre pago y resultados). El resto del capítulo puede verse como un intento del campo de la Economía de las Organizaciones de remediar estas limitaciones y acercar la teoría a la práctica de la provisión de incentivos en las organizaciones.

3. Distorsiones en la provisión de incentivos: la insensatez de pagar por A esperando que se haga B

El modelo clásico de agencia de la sección 2 señala al riesgo como la consideración central en la provisión de incentivos. Sin embargo, en muchos contratos del mundo real, el riesgo no parece ser un factor crucial: como argumenta Baker (2002), con la excepción de los planes de pago en acciones para los ejecutivos de alto rango, la mayoría de los acuerdos de compensación imponen poco riesgo a los empleados. Los datos tampoco parecen reflejar la relación negativa entre riesgo e incentivos que sugiere (2) – ver Prendergast (2000).

Si bien existe cierta evidencia empírica de que la disyuntiva entre riesgo e incentivos tiene algún poder para explicar la forma de los contratos, también es cierto que la misma deja sin explicar buena parte de la variabilidad contractual (ver Gibbons, 1998, para citas al respecto). La literatura reciente de *multitasking* (ver Holmström y Milgrom, 1991, y Baker, 1992, para las contribuciones originales) enfatiza una segunda consideración en el diseño de un contrato de incen-

tivos: la distorsión que el mismo podría inducir en el comportamiento de un agente que tiene a su cargo múltiples tareas cuando el principal, ante la imposibilidad de basar el contrato en lo que realmente le importa (digamos, el cambio en el valor de la empresa – B), recurre a una medida alternativa del desempeño del agente (por ejemplo, la ganancia contable – A).

En términos generales, podemos pensar a y como el aporte que el agente hace al objetivo del principal. El llamarlo ‘output’ como hasta ahora podría inducir a pensar que este aporte individual es fácilmente medible. De hecho, el modelo clásico implícitamente hace este supuesto cuando permite que los contratos estén atados a y . Consideremos el caso en el que y denota el valor total de la empresa. Si la empresa cotiza en Bolsa, el valor de la empresa está bien aproximado por el valor de mercado de sus acciones; en ese caso, suponer que y puede incluirse en un contrato parece razonable y, de hecho, el valor de las acciones es comúnmente utilizado en los paquetes de compensación de los ejecutivos. El uso de y , por definición, provee incentivos que no generan ninguna distorsión en el comportamiento.

¿Pero qué ocurre si la empresa no cotiza en Bolsa? Claramente el contrato no puede basarse en el valor de mercado de las acciones. Aun si la empresa cotizara en Bolsa, el valor de las acciones podría resultar inútil como medida del desempeño para empleados en niveles más bajos de la jerarquía – típicamente porque es una medida muy ruidosa de la contribución del agente o, dicho de otro modo, porque el agente tiene poco control sobre el valor de la empresa a través de sus acciones. En el marco del modelo clásico, tendríamos $\sigma^2 \rightarrow \infty$, lo que implicaría $b \rightarrow 0$ aun si $r \approx 0$ – ver (2).

Claramente, en cualquiera de los casos la empresa no puede basar sus contratos en y y debe buscar medidas alter-

nativas del desempeño. Para continuar con el ejemplo, consideremos una empresa que decide remunerar a sus gerentes sobre la base de la ganancia contable. ¿Es esto una buena idea? Supongamos, como en Baker (2002), que el agente tiene que decidir si invertir en un proyecto de I+D. Tal inversión podría tener un impacto positivo sobre el precio de la acción, pero afectar negativamente la ganancia contable del año (pues los costos se pagan en el ejercicio, pero los beneficios potenciales sólo se perciben en plazos más largos). Un agente que es remunerado de acuerdo a la ganancia contable dejaría pasar esta oportunidad de contribuir al valor de la empresa. Es más: el agente así remunerado tendría incentivos a manipular la contabilidad de la empresa; ello no afecta el precio de la acción (si el mercado percibe la manipulación), pero podría aumentar la ganancia contable y, por ende, la compensación del agente.¹⁶

Para modelar estos fenómenos de distorsión, supongamos entonces que y es una medida de la contribución que el agente hace a la organización (o a lo que el principal valora) y que no puede incluirse en un contrato. En principio puede ser observable, pero supondremos que no es verificable por un tercero: los contratos basados en y no pueden hacerse valer ante un juez. Suponemos también que existe una medida alternativa del desempeño, que llamaremos p , y que sí podría ser incluida en un contrato formal.¹⁷ Si $y = a + \varepsilon$ y $p = a + \phi$, un contrato $w(p) = s + b \cdot p$ crea incentivos a incrementar el valor de p , pero esas acciones también mejoran y . En ese caso, el uso de una medida alternativa no genera nin-

¹⁶ Existe abundante evidencia sobre las manipulaciones inducidas por los esquemas de compensación; ver, por ejemplo, Courty y Marschke (2004) y Oyer (1998).

¹⁷ En términos generales, esta medida alternativa suele ser una representación incompleta o imperfecta de las consecuencias económicas de las acciones del agente.

guna distorsión. Pero supongamos, en cambio, que hay dos acciones o que la acción tiene dos dimensiones, que llamamos a_1 y a_2 . Si $y = a_1 + \varepsilon$ y $p = a_2 + \phi$, un contrato basado en p no puede generar ningún valor para el principal (quien lógicamente elegiría $b = 0$). Si bien extremo, el ejemplo ilustra la distorsión inducida por el uso de una medida alternativa de desempeño y nos introduce al mensaje principal de estos modelos: no tiene sentido generar incentivos poderosos para que el agente lleve a cabo las acciones incorrectas.

Para formalizar estas ideas, recurriremos una vez más a Gibbons (2005b y 2010b). Como en el caso del modelo clásico, presentamos primero la cronología de eventos:

1. El principal y el agente firman un contrato que especifica la compensación del agente como una función del output. Nos limitaremos a considerar contratos en los que la compensación total es una función lineal de la medida de desempeño:

$$w(p) = s + b \cdot p,$$

donde s y b tienen la misma interpretación que antes.

2. El agente elige las acciones a_1 y a_2 , que el principal no observa. Los costos personales que enfrenta el agente vienen dados por $c(a_1, a_2) = \frac{1}{2}a_1^2 + \frac{1}{2}a_2^2$, con $c(0,0) = 0$.
3. Tienen lugar ciertos eventos fuera del control del agente (ε y ϕ) que afectan tanto el output como la medida de desempeño.
4. Las acciones escogidas y los ruidos determinan la contribución del agente al valor de la empresa (y) y la performance medida (p). La tecnología de producción es $y = f_1a_1 + f_2a_2 + \varepsilon$, la tecnología de medición del desempeño es $p = g_1a_1 + g_2a_2 + \phi$, y los

ruidos ε y ϕ son variables aleatorias independientes con media cero.¹⁸

5. El principal observa la *performance* y el agente recibe la compensación especificada en el contrato, por lo que su compensación neta es $w(p) - c(a_1, a_2)$. El principal se queda con el output, pero debe pagar la compensación del agente, con lo que su pago neto es $y - w(p)$. Tanto el principal como el agente son neutrales al riesgo, por lo que simplemente buscan maximizar el valor esperado de estos pagos.

Las acciones de *first best* son las que resultan de maximizar el excedente total $E[y - c(a_1, a_2)]$ y vienen dadas por:

$$a_i^{FB} = f_i, \quad i = 1, 2.$$

El problema del agente es:

$$\max_{a_1, a_2} E[w(p) - c(a_1, a_2)] = s + b(g_1 a_1 + g_2 a_2) - \frac{1}{2} a_1^2 - \frac{1}{2} a_2^2,$$

cuya solución $a_i(b) = g_i \cdot b$ nos da las acciones óptimas para el agente dado el contrato $w(p)$. El problema del principal puede escribirse, de manera análoga a la sección 2, como:

$$\max_b f_1 a_1(b) + f_2 a_2(b) - \frac{1}{2} [a_1(b)]^2 - \frac{1}{2} [a_2(b)]^2 = f_1 g_1 b + f_2 g_2 b - \frac{1}{2} g_1^2 b^2 - \frac{1}{2} g_2^2 b^2.$$

De la condición de primer orden de este problema obtenemos el valor eficiente de b :

¹⁸ Dado el supuesto de neutralidad al riesgo del agente, no nos preocupamos por la varianza de los términos de error.

$$(3) \quad b^* = \frac{f_1 g_1 + f_2 g_2}{g_1^2 + g_2^2} = \frac{\|f\|}{\|g\|} \cos(\theta),$$

donde $\|x\| = \sqrt{x_1^2 + x_2^2}$ es la norma del vector x y θ es el ángulo entre los vectores $f = (f_1, f_2)$ y $g = (g_1, g_2)$.¹⁹ El primer componente de (3), $\|f\| / \|g\|$, representa un ajuste de escala. Si g_1 y g_2 fueran mucho mayores que f_1 y f_2 , unas acciones dadas del agente tendrían un gran efecto sobre p , y por ende sobre $w(p)$, pero un pequeño impacto sobre lo que al principal realmente le interesa, y . Lógicamente, el principal querría reducir b en este caso; eso es exactamente lo que prescribe (3), ya que $\|f\| / \|g\|$ sería pequeño en este caso.

Más importante es el segundo factor en (3), $\cos(\theta)$, que mide la distorsión en una medida de desempeño. Para entenderlo, conviene comenzar por los casos extremos. Supongamos entonces que $y = a_1 + \varepsilon$ y $p = a_1 + \phi$; por lo tanto, $f = g = (1, 0)$ – los vectores están perfectamente alineados y un contrato basado en p reproduce exactamente los incentivos que brindaría un contrato basado en y (que, por definición, no tiene distorsión). En este caso, $\theta = 0^\circ$ y $\cos(\theta) = 1$. Por lo tanto, $b^* = 1$, $a_1^* = a_1^{FB} = 1$ y $a_2^* = a_2^{FB} = 0$ (claramente, la segunda acción no tiene ningún valor para el principal).

Supongamos, por el contrario, que $y = a_1 + \varepsilon$ y $p = a_2 + \phi$; por lo tanto, $f = (1, 0)$ y $g = (0, 1)$ – los vectores son ortogonales y un contrato basado en p no puede generar ningún valor para el principal. El producto marginal de la acción 2 sobre el valor de la empresa es cero, por lo que el agente puede incrementar su compensación sin incrementar el valor de la organización. Como $\theta = 90^\circ$ en este caso, $\cos(\theta) = 0$ y, como era de esperarse, $b^* = 0$ (no tiene sentido pagar por las

¹⁹ La segunda parte de (3) aprovecha la relación existente entre el ángulo θ entre dos vectores y el producto interior de los mismos: $f \cdot g = \|f\| \|g\| \cos(\theta)$.

acciones incorrectas).²⁰ La regla general a extraer de estos ejemplos es que cuanto menor sea $\cos(\theta)$, mayor será la distorsión inducida por la medida alternativa de desempeño y menores los incentivos brindados por el contrato óptimo.²¹ $\cos(\theta)$, por lo tanto, resume qué tanto la medida de desempeño genera señales apropiadas para el agente respecto de qué acciones emprender (Baker, 2002).

La expresión para el contrato óptimo, (3), deja en claro que el supuesto de neutralidad al riesgo (o la existencia de una medida de performance sin ruido) no es suficiente para lograr incentivos perfectos. Además, la medida de desempeño debe ser perfectamente congruente con el objetivo del principal – es decir, $\theta = 0^\circ$. Nótese que la existencia de ruido no impide el logro del *first best*; simplemente lo hace no rentable en el caso de un agente averso al riesgo (debido a la prima que hay que pagarle). La no congruencia de la medida de desempeño, por el contrario, vuelve imposible alcanzar el *first best* (Feltham y Xie, 1994).

No todas las medidas alternativas del desempeño (incluyendo lo que el principal valora) tienen la misma sensibilidad a todas las dimensiones de las acciones emprendidas por el agente. Los problemas con una medida de desempeño dada aparecen cuando los productos marginales de las acciones sobre la medida de desempeño (g) difieren de los productos marginales de las acciones sobre el valor de la empresa (f). Esta distinción es clave: lo que caracteriza a una buena medida de desempeño es la correlación entre estos

²⁰ Lincoln Electric es un caso paradigmático en el uso extensivo de incentivos (ver Fast y Berg, 1975). En un momento, la empresa pagaba a sus secretarías de acuerdo al número de letras tecleadas. La práctica fue abandonada cuando se descubrió que las secretarías pasaban su hora de almuerzo pulsando sin parar la misma tecla (Baker, 2002).

²¹ Feltham y Xie (1994) desarrollan una medida alternativa de la congruencia: $\delta = [f_1g_2 - f_2g_1]^2$. En ese caso, $\delta = 0$ si, y sólo si, existe $\lambda \neq 0$ tal que $g_i = \lambda g_i -$ lo cual es equivalente a $\theta = 0^\circ$.

márgenes (tal y como es capturada por $\cos(\theta)$), y no la correlación entre y y p . Si $y = a_1 + \varepsilon$ y $p = a_2 + \varepsilon$, la correlación entre y y p es alta, pero claramente p es una pésima medida del desempeño. Retomando el ejemplo anterior, lo que importa no es si la ganancia contable está correlacionada con el valor de la empresa o el precio de la acción (lo cual probablemente sea cierto), sino si la ganancia contable responde a las acciones del agente de la misma manera que el precio de la acción. El principal podría otorgarle un peso relativamente bajo a la ganancia contable en el contrato óptimo, aun si la misma está muy correlacionada con el precio de la acción, si darle demasiada importancia a la ganancia contable generara comportamientos disfuncionales (Baker, 2002).

En el diseño de incentivos ligados al desempeño, el camino al infierno está plagado de buenas intenciones, como ilustra la siguiente cita:

“Business history is littered with firms that got what they paid for. At the H. J. Heinz Company, for example, division managers received bonuses only if earnings increased from the prior year. The managers delivered consistent earnings growth by manipulating the timing of shipments to customers and by pre-paying for services not yet received [...]. At Dun & Bradstreet, salespeople earned no commission unless the customer bought a larger subscription to the firm's credit-report services than in the previous year. In 1989 the company faced millions of dollars in lawsuits following charges that its salespeople deceived customers into buying larger subscriptions by fraudulently overstating their historical usage [...]. In 1992 Sears abolished the commission plan in its auto-repair shops, which paid mechanics based on the profits from repairs authorized by customers. Mechanics misled customers into authorizing unnecessary repairs, leading California officials to prepare to close Sears' auto-repair business statewide [...].”

Baker et al. (1994)

En todos estos casos, en los que el principal no podía pagar sobre la base de lo que realmente le importa, el uso de una medida alternativa del desempeño condujo a distorsiones en el comportamiento de los empleados, quienes buscando maximizar su compensación emprendieron acciones dañinas para el valor de la empresa.

Riesgo y distorsión

En el caso más general (es decir, sin el supuesto de neutralidad al riesgo que hicimos aquí a fines expositivos), aparecen dos consideraciones en el uso de cualquier medida de desempeño: riesgo y distorsión. Si la empresa tiene a su alcance una medida contratable del desempeño que combine bajo ruido y baja distorsión, puede utilizar incentivos fuertes (Baker, 2002). Es el caso del fabricante de parabrisas Safelite Glass Corporation (Lazear, 2000): entre 1994 y 1995, la empresa pasó gradualmente de un mecanismo de compensación por hora trabajada a otro por parabrisas instalado y, en poco más de un año, la producción por trabajador aumentó un 44%, del cual la mitad puede atribuirse a un incremento en el esfuerzo de los empleados en respuesta a los mayores incentivos. Lamentablemente, las empresas no siempre tienen medidas tan buenas del desempeño o la contribución individual (Gibbons, 1998); en ese caso, será óptimo brindar incentivos débiles.

En la elección de una medida de desempeño, las empresas típicamente se enfrentan a una disyuntiva entre medidas con alto riesgo y baja distorsión, y medidas con bajo riesgo y alta distorsión. Tomemos el caso del diseño de incentivos para gerentes de división. En este ejemplo, la interacción entre divisiones se vuelve una consideración crucial: si la interacción es importante, las medidas que minimizan la distorsión se verán favorecidas aunque sean más riesgosas (por ejemplo, las ganancias globales de la compañía); si la

interacción es mínima, poco se pierde en términos de distorsión por escoger una medida menos riesgosa o más controlable para el gerente (como las ganancias de su división).²² Como muestran Datar et al. (2001), en general, el ruido en las medidas de desempeño fuerza al principal a sacrificar congruencia (y tolerar distorsión) en aras de reducir el riesgo en la compensación del agente y, por ende, la prima por riesgo que debe pagarle.

Múltiples medidas de performance

Hasta aquí supusimos que el principal tenía a su disposición una única medida de desempeño p para incluir en el contrato. Cuando existen más medidas alternativas, aparecen nuevas preguntas interesantes como, por ejemplo, si es valioso tener medidas adicionales o cómo se las debe combinar de manera óptima en un contrato de incentivos.

Incrementar el número de medidas de desempeño podría reducir el riesgo al que es sujeto el agente o proveer una mejor estimación del desempeño que facilite la inducción de acciones más congruentes con los objetivos del principal (es decir que incrementar el número de medidas puede agrandar el conjunto de acciones implementables – hasta, potencialmente, incluir el *first best*). Como muestran Feltham y Xie (1994), el valor de una medida adicional es cero si, y sólo si, las medidas existentes constituyen un estadístico suficiente para la medida adicional respecto de las acciones del agente.

Dejando de lado, una vez más, al riesgo, podemos modificar el modelo anterior para considerar el caso de múltiples medidas de desempeño. Supongamos entonces que el principal tiene a su alcance dos medidas, $p_1 = g_{11}a_1 + g_{12}a_2 + \phi_1$ y $p_2 = g_{21}a_1 + g_{22}a_2 + \phi_2$. La respuesta óptima del agente al contrato lineal $w(p_1, p_2) = s + b_1p_1 + b_2p_2$ viene dada por:

²² El ejemplo es de Baker (2002).

$$a_i(b_1, b_2) = b_1 \cdot g_{1i} + b_2 \cdot g_{2i}$$

y el problema del principal puede ahora escribirse como:

$$\max_b f_1 a_1(b_1, b_2) + f_2 a_2(b_1, b_2) - \frac{1}{2} [a_1(b_1, b_2)]^2 - \frac{1}{2} [a_2(b_1, b_2)]^2$$

o:

$$(4) \quad \min_{b_1, b_2} \vartheta \equiv \sum_{j=1}^2 \left(f_j - \sum_{i=1}^2 b_i g_{ij} \right)^2.$$

La segunda representación del problema del principal pone de manifiesto que el mismo escoge los ponderadores de las medidas de desempeño en el contrato óptimo (b_1 y b_2) de manera de minimizar la incongruencia entre la compensación global del agente y el objetivo del principal. Sea $d = (b_1 g_{11} + b_2 g_{21}, b_2 g_{12} + b_2 g_{22})$. Minimizar ϑ , de hecho, es equivalente a minimizar la distancia (o el ángulo) entre f y d , el vector que muestra cómo responde la compensación del agente a sus acciones, dado el contrato. La congruencia es endógena en este caso, ya que depende de cómo se ponderen las medidas de desempeño. La solución al problema (4) es:

$$(5) \quad b_i^* = \frac{f_i g_{jj} - f_j g_{ji}}{g_{11} g_{22} - g_{12} g_{21}}.$$

Las acciones de *first best* no dependen del número de medidas disponibles, por lo que siguen siendo $a_i^{FB} = f_i$, $i = 1, 2$. Para alcanzar el resultado de *first best* debe verificarse entonces que:

$$(6) \quad f_i = b_1 \cdot g_{1i} + b_2 \cdot g_{2i} \iff f_i - \sum_{j=1}^2 b_j g_{ji} = 0,$$

o, en otras palabras, el *first best* es alcanzable si, y sólo si, el contrato óptimo genera una medida global del desempeño perfectamente congruente. Reemplazando (5) en (6), puede chequearse que la condición se cumple en este caso: con sólo dos acciones, pueden brindarse incentivos perfectos siempre y cuando los dos vectores de productos marginales de las acciones, g_1 y g_2 , no sean perfectamente colineales (Datar et al., 2001). Si hubiera más de dos acciones, sería más difícil alcanzar la congruencia perfecta, porque el principal tendría menos instrumentos que acciones a controlar – éste, de hecho, era el problema cuando discutimos el caso de una única medida de desempeño. En líneas generales, el *first best* puede alcanzarse, bajo neutralidad al riesgo, con múltiples medidas siempre que las mismas abarquen la dirección de *first best* (es decir, la del vector f) – ver Feltham y Xie (1994).²³

En el caso general con un agente con aversión al riesgo, Feltham y Xie (1994) demuestran que el *first best* puede alcanzarse con una única medida de desempeño si, y sólo si, la medida es perfectamente congruente y sin ruido (es decir, completamente bajo el control del agente).²⁴ Datar et al. (2001) extienden esta noción al caso de múltiples medidas de desempeño y muestran que los ponderadores en el contrato óptimo deben escogerse para minimizar la suma de la incongruencia y la prima de riesgo a pagar al agente.

²³ En general, n vectores linealmente independientes forman una base de \mathfrak{R}^n . Por lo tanto, cualquier vector $f \in \mathfrak{R}^n$ puede ser replicado eligiendo óptimamente los ponderadores del contrato cuando el principal cuenta con al menos n medidas de desempeño (no degeneradas).

²⁴ La segunda condición se volvía irrelevante bajo el supuesto de neutralidad al riesgo.

Instrumentos alternativos

La predicción básica de la literatura de *multitasking* es que cuando las medidas alternativas contienen mucho ruido y generan mucha distorsión, los incentivos serán óptimamente débiles. Es en estos casos que la búsqueda de instrumentos alternativos se vuelve aún más acuciante.

Holmström y Milgrom (1991) exploran el caso de una empresa que debe decidir si imponer restricciones a las actividades comerciales personales durante el horario laboral. Como explican estos autores, no son tanto las características de estas actividades lo que determina si deberían ser permitidas, sino cómo interactúan con otras actividades que debe emprender el agente. La idea general es que hay dos formas de incentivar una tarea particular: (i) recompensar la tarea en sí y (ii) reducir el costo de oportunidad de la misma eliminando los incentivos a realizar otras tareas que compiten por el tiempo y la atención del agente.

Para ilustrar el rol de las limitaciones sobre las actividades externas del agente [(ii)], supongamos que y no es contractable y que $y = a_1$, pero que el principal tiene a su disposición una medida de desempeño $p = a_1 + a_2$. Un contrato basado en p genera incentivos para realizar a_2 , aunque esa acción no contribuye en nada al objetivo del principal. Sabemos de (3) que $b^* = 1/2 < b^{FB} = 1$ en este caso – los incentivos monetarios son menos que perfectos. Pero supongamos ahora que el principal puede prohibir la segunda tarea ($a_2^* = 0$). En ese caso, la performance medida es $p = a_1$ y el principal puede inducir el *first best* ($a_1^* = a_1^{FB} = 1$) fijando $b^* = 1$.

Otro instrumento alternativo estudiado en Holmström y Milgrom (1991) es la propiedad de los activos. Los autores utilizan esta variable para distinguir entre empleados y contratistas independientes: los primeros utilizan activos que son propiedad de otros, mientras que los segundos utilizan activos propios. Para modelar de manera simple este caso,

recurrimos a Gibbons (2010a) y agregamos al modelo de la sección 3 una máquina que el agente utiliza para producir y . El valor de reventa de esta máquina es $v = h_1 a_1 + h_2 a_2 + \xi$, que suponemos no contractable, como y , por lo que los contratos sólo pueden depender de p . Quien tiene la propiedad de la máquina, recibe v al final del período. En cualquier caso, el *first best* requiere ahora escoger $a_i^{FB} = f_i + h_i$.

El principal tiene en esta situación dos instrumentos para afectar los incentivos del agente y , generalmente, el uso de ambos será optimizado conjuntamente – es decir, el contrato formal basado en p típicamente será diferente de acuerdo a quién tenga la propiedad del activo. Consideremos un ejemplo simple en el que $y = a_1$, $v = a_2$ y $p = a_1$. Un contrato basado en p no puede, por sí solo, generar incentivos correctos para incrementar v , pero es fácil ver que en combinación con la propiedad del activo en manos del agente, permite alcanzar el *first best*. Si, por el contrario, $y = a_1$, $v = a_2$ y $p = a_1 + a_2$, el mismo arreglo (el agente es dueño de la máquina) genera demasiados incentivos a incrementar v (es decir, a realizar la segunda tarea). En este caso, el *first best* es alcanzable si el principal retiene la propiedad del activo y ofrece un contrato basado en p , con $b^* = 1$.

Si bien son extremadamente simples, estos ejemplos resaltan dos ideas fundamentales: (i) es importante estudiar todo el rango de las actividades del agente para analizar sus incentivos y (ii) el uso de múltiples instrumentos puede ayudar a brindar un conjunto más balanceado de incentivos. En la sección 5 vamos a elaborar la segunda idea y a hacerla más precisa, pero primero analizaremos en la sección siguiente un instrumento más: las evaluaciones subjetivas del desempeño.

4. Contratos informales: provisión de incentivos (casi) sin contrato

La teoría estándar de agencia de la sección 2 tiene problemas para explicar la prevalencia de contratos que simplemente especifican un salario fijo y, en general, para justificar la debilidad de los incentivos brindados dentro de las empresas – especialmente, comparados con los que brinda el mercado (Holmström y Milgrom, 1991).²⁵ La teoría de *multitasking* de la sección 3 sugiere, por el contrario, que esta relativa ausencia de incentivos explícitos puede ser generalizada, debido a consideraciones de riesgo y, fundamentalmente, distorsión. Si bien otros instrumentos formales alternativos (como la propiedad de los activos y las limitaciones sobre las actividades permisibles) pueden ayudar a generar incentivos más fuertes, típicamente los contratos óptimos brindan incentivos débiles en muchos casos. A pesar de ello, solemos encontrar altos niveles de esfuerzo en situaciones en las que no hay un vínculo contractual entre compensación y resultados. De hecho, en muchas circunstancias, las partes parecen arreglarse muy bien (casi) sin contrato:

“...if something comes up, you get the other man on the telephone and deal with the problem. You don't read legalistic contract clauses at each other if you ever want to do business again. One doesn't run to lawyers if he wants to stay in business because one must behave decently”.

“You can settle any dispute if you keep the lawyers and accountants out of it. They just do not understand the give-and-take needed in business”.

Citados en Stewart Macaulay (1963)

²⁵ Para una revisión de la evidencia, consultar MacDonald (1984), Hart y Holmström (1987) y Baker et al. (1988).

Las citas ilustran la idea de que las partes pueden usar su subjetividad o discrecionalidad para acomodar los vaivenes de la relación y sugieren que la interacción repetida en el tiempo (“if you ever want to do business again”) juega un rol importante en la provisión de incentivos.

Como resultado de las limitaciones de los instrumentos formales, los incentivos en las organizaciones suelen ser informales (Levin, 2003). Por ejemplo, la mayor parte de los empleados de una empresa no tiene un pago contractualmente ligado a su desempeño, pero las evaluaciones “subjetivas” son importantes para determinar el salario futuro y las decisiones sobre ascensos y despidos. Estas decisiones suelen estar basadas en criterios como capacidad para el trabajo en equipo, liderazgo o iniciativa – todas dimensiones difíciles de verificar y, por lo tanto, de incluir en un contrato, aun cuando el desempeño del empleado en todas estas dimensiones sea claro para sus superiores.²⁶

Al no estar asentados formalmente, estos contratos implícitos (o relacionales) son sólo promesas informales; por lo tanto, debe estar en el interés de las partes respetarlos o, dicho de otro modo, el valor futuro de la relación debe ser suficientemente grande como para que ninguna parte falte a su palabra. La ventaja del contrato informal radica en que, mientras que un contrato formal (como los de las secciones precedentes) debe ser especificado ex ante en términos que puedan ser verificados ex post por un tercero (por ejemplo, un juez), un contrato informal puede estar basado en resultados que sólo las partes pueden observar ex post, o que son prohibitivamente costosos de especificar ex ante (Baker et al., 2002).

²⁶ Queda claro, entonces, que la relativa incapacidad para observar las acciones no es la única fuente de problemas de motivación. Aun si el comportamiento es observado por todos, si el mismo es difícil de verificar para una tercera parte los contratos que prescriben acciones determinadas rara vez serán efectivos para motivar (Roberts, 2004).

Estos contratos informales son ubicuos en la vida de las organizaciones y afectan el comportamiento de sus miembros de manera sustancial.²⁷ Muchos procesos formales, de hecho, no pueden ser comprendidos sin tener en cuenta los acuerdos informales asociados a los mismos:

“The distinction between the ‘formal’ and ‘informal’ organization of the firm is one of the oldest in the literature, and it hardly needs repeating that observers who assume firms to be structured in fact by the official organization chart are sociological babes in the woods”.

Mark Granovetter (1985)

Para reforzar el punto, consideremos el caso de United Airlines (relatado en Levin, 2003). En el verano boreal de 2001, durante las negociaciones salariales, el sindicato de pilotos dispuso que sus miembros trabajaran “a reglamento”. La consecuencia de esta medida (de respetar al pie de la letra lo escrito en el contrato formal) fue un verano de demoras y cancelaciones de vuelos que le costaron 700 millones de dólares a la empresa. Claramente, entender los aspectos informales en la organización y su interacción con los aspectos formales es crucial.

Para formalizar estos contratos informales, construimos un modelo simple basado en Gibbons (2005b). Un principal y un agente interactúan repetidamente. La cronología de eventos en cada período de la relación entre el principal y el agente es la siguiente:

1. El principal ofrece al agente un acuerdo de compensación (s, b) que especifica un salario básico s y promete pagar un bonus b si el output del agente es alto ($y = H$, ver abajo). La compensación del agente, por lo tanto, es:

²⁷ Ver Baker et al. (2002) para una discusión más extensa.

$$w(y) = \begin{cases} s + b & \text{si } y = H \\ s & \text{si } y = L \end{cases} .$$

2. Si el agente rechaza la oferta, toma una oportunidad laboral alternativa que le ofrece un pago w_a . Si acepta el acuerdo, el agente elige una acción $a \in [0, 1]$, que el principal no observa, a un costo $c(a) = a^2/2$, con $c(0) = 0$.
3. El principal y el agente observan la realización del output y . Sin embargo, y no puede ser medido objetivamente ni verificado por terceras partes.²⁸ Supondremos que y puede ser alto ($y = H$) o bajo ($y = L < H$), y que la acción escogida por el agente determina la probabilidad de tener un resultado bueno: $Pr(y = H | a) = a$.
4. El agente recibe el salario especificado en el contrato y , si $y = H$, el principal decide si paga o no el bonus prometido. La compensación neta del agente es $w - c(a)$ y la del principal, $y - w$. Como siempre, hacemos el supuesto de que tanto el principal como el agente son neutrales al riesgo, por lo que simplemente buscan maximizar el valor esperado de estos pagos.

El *first best* es el resultado de escoger la acción que maximiza el excedente total $E[y - c(a)] = L + a(H - L) - a^2/2$. Por lo tanto, $a^{FB} = H - L$. Supongamos, para comenzar, que la relación dura un único período y resolvamos el juego hacia atrás. Como las partes no vuelven a hacer negocio juntas en el futuro, el principal no tiene ningún incentivo a pagar b en $t = 4$. Anticipando esto, el agente escoge $a = 0$ en $t = 2$, por lo que $y = L$ con probabilidad 1. Si suponemos que

²⁸ Suele decirse que y es observable, pero no verificable, en este caso.

$L < w_a$, el principal no puede ofrecer ningún salario s que el agente aceptaría y, por ende, la relación nunca se concreta. En este equilibrio estático, las ganancias del intercambio no se realizan y ambas partes reciben los pagos correspondientes a su mejor alternativa (w_a para el agente y, por simplicidad, 0 para el principal).

¿Puede el principal lograr un resultado más satisfactorio en una relación más duradera? Típicamente, sí. En una relación continuada, las partes pueden basar los términos de la interacción futura en el éxito de la interacción presente, lo que potencialmente permite el logro de un acuerdo que provea incentivos efectivos (Levin, 2003). Supongamos, entonces, que el juego entre el principal y el agente se repite indefinidamente, y que las partes descuentan el futuro a una tasa r . Nos concentramos en estrategias “gatillo” (*trigger strategies*) en las que las partes comienzan cooperando y luego mantienen la cooperación si, y sólo si, la otra parte no faltó a su palabra en el período anterior – en cuyo caso la parte ofendida se niega a cooperar en el futuro. Analizamos a continuación bajo qué condiciones estas estrategias resultan en un equilibrio del juego repetido y nos preguntamos si el resultado obtenido es mejor que en el caso de un período.

Si el agente piensa que el principal va a cumplir en $t = 4$ con su promesa de pagar b si $y = H$, en $t = 2$ resuelve el siguiente problema:

$$\max_{\alpha \in [0,1]} s + b \cdot \alpha - \frac{1}{2}\alpha^2,$$

cuya solución es:

$$(7) \quad a(b) = b$$

Si el principal ofrece el mínimo pago esperado que induce al agente a aceptar el contrato, $s + b \cdot a(b) = \frac{1}{2}[a(b)]^2 + w_a$, podemos escribir el pago esperado del principal en cada período como

$$V(b) \equiv L + a(b)(H - L) - \left\{ \frac{1}{2} [a(b)]^2 + w_a \right\} = \text{output esperado} - \text{compensación del agente.}$$

Como el pago contingente en el output es meramente una promesa, existe la tentación para el principal de renegar del acuerdo una vez que el output se ha realizado. La pregunta relevante ahora es si el agente debe creer que el principal va a pagar el bonus o, más directamente, si al principal le conviene pagar b en caso que $y = H$. Para responder a esta pregunta, comparemos las alternativas del principal. Si el principal honra su promesa recibe $H - s - b$ en el período, y $V(b)$ en cada período de allí en más. Su pago en esta opción es

$$H - s - b + V(b)/r.$$

Si, por el contrario, el principal no pagara el bonus prometido, recibiría $H - s$ en el período, pero el agente no cooperaría nunca más con él y el pago en todo período futuro sería 0.²⁹ Por lo tanto, estará en el interés del principal respetar el contrato implícito siempre que

$$H - s - b + V(b)/r \geq H - s,$$

que puede reescribirse como

$$(8) \quad b \leq V(b)/r.$$

²⁹ Formalmente, las partes revierten al equilibrio de Nash del juego de un período que analizamos anteriormente.

El contrato óptimo maximiza $V(b)$ sujeto a la restricción (8). Para lograr el *first best*, el bonus que debe prometer el principal es $b^{FB} = H - L$; chequear (7). Tal promesa sólo será creíble si el valor futuro de la relación, medido por $V(b)/r$, es lo suficientemente grande. De la inspección de la condición (8) se desprende que el *first best* será implementable si la tasa r es lo suficientemente baja; es decir, si el principal valora el futuro lo suficiente o es suficientemente paciente. Pero si la tasa de interés es más alta, el valor presente de la interacción con el agente no es tan alto como para hacer creíble un bonus tan grande, y el principal sólo puede prometer creíblemente un bonus menor: el agente tendrá menos incentivos y el valor esperado de la relación caerá, pero el principal aún tendrá incentivos a cumplir su promesa si el agente realiza una contribución alta. Finalmente, si r es suficientemente alta no hay ninguna promesa que sea creíble, por lo que no habrá ni bonus ni incentivos. En cualquier otro caso, el contrato implícito permite a las partes mejorar su situación con respecto al caso en que no interactuaban reiteradamente.

La condición (8) ilustra la idea general de que, en un contrato implícito, hay un límite a lo que puede ser prometido de manera creíble. El principio general es que una promesa dada será creíble sólo si la máxima tentación a violar el acuerdo (en nuestro caso es b , lo que el principal puede ahorrarse incumpliendo cuando $y = H$) es menor que el valor presente del excedente neto generado por la relación. Baker et al. (1994) estudian el uso conjunto de contratos formales (basados en p) y contratos informales (basados en y), y muestran que el contrato informal puede reducir las distorsiones causadas por el contrato formal, mientras que el contrato formal puede reducir la tentación del principal a romper sus promesas (y, por ende, resultar en un bonus menor para iguales incentivos). El tema general aquí es, una vez

más, que los aspectos formales e informales no sólo coexisten, sino que también interactúan, por lo que deben diseñarse conjuntamente.

El rol del manager

Una idea central de esta sección es que un contrato informal puede ayudar a las partes a realizar las ganancias del intercambio en circunstancias en las que los instrumentos formales fallan. Esto pone de relieve un rol potencialmente importante de los managers: el desarrollo y mantenimiento de estos contratos informales dentro de la empresa (Baker et al., 2002). En el modelo precedente, este rol se reduce a definir las medidas del desempeño a considerar (y), el nivel de desempeño que amerita el bonus ($y = H$) y el monto del bonus (b), así como comunicar todas estas decisiones al agente, verificar los resultados y decidir si vale la pena honrar el contrato informal (de acuerdo a (8)).

En el modelo simple que hemos bosquejado, estas tareas son triviales, aunque distan mucho de serlo en la práctica. Más aún, nos hemos limitado a considerar exclusivamente la credibilidad de las promesas, dejando de lado potenciales problemas en la claridad con que las mismas son formuladas y transmitidas (Gibbons, 2010a). Finalmente, más difícil que diseñar y mantener un contrato informal parece ser cambiarlo por otro: ello involucra romper una promesa, al tiempo que se pretende ser creíble al formular una nueva (Gibbons, 2006).

En estos márgenes de la organización interna de la empresa se juega una parte de la creación de una ventaja competitiva – la idea de que empresas similares pueden tener desempeños persistentemente diferentes porque implementaron diferentes prácticas de gestión (incluyendo contratos informales) que no son fáciles de imitar. Consultar Gibbons (2010a) para estudios de caso sobre la construcción y cambio de contratos informales, así como embriones de nuevas teorías al respecto.

5. Sistemas de incentivos

El problema de motivación en las organizaciones que hemos abordado en este capítulo consiste en lograr que los intereses de los miembros de la organización estén lo más alineados posible con los intereses de la organización misma (Roberts, 2004). En esta tarea, los responsables de la organización tienen a su alcance muchos más instrumentos que los contratos formales enfatizados por la teoría de agencia clásica o la más reciente de *multitasking*. De hecho, todos los aspectos organizacionales pueden ser empleados; en este capítulo discutimos algunos de ellos (como la propiedad de los activos, las limitaciones en las tareas que puede emprender un empleado y los contratos informales), pero la lista es mucho más extensa: reparto de tareas, selección de personal, decisiones de ascensos y despidos, organización jerárquica de la empresa, delegación de autoridad, comunicación, ambiente de trabajo, relaciones laborales, cultura corporativa... y la lista sigue.

Una idea central que los modelos presentados en este capítulo intentaban transmitir es que ningún instrumento aislado suele ser capaz de generar incentivos perfectos o, al menos, satisfactorios, por lo que es razonable utilizar múltiples instrumentos para resolver el problema de motivación, en un verdadero *sistema de incentivos* (Holmström y Milgrom, 1994). La existencia de relaciones entre los distintos instrumentos tiene como consecuencia la aparición de un número restringido de patrones coherentes, en los cuales los distintos instrumentos “encajan” (*fit*) bien entre sí y con el entorno en el que opera la empresa y, por lo tanto, producen una buena performance. El problema de la organización puede interpretarse también como el de identificar y seleccionar el mejor patrón dentro de este conjunto (Roberts, 2004).

La noción clave detrás de estos patrones coherentes es la de complementariedad: dos instrumentos son complementarios cuando utilizar (más) un instrumento incrementa el ren-

dimiento de utilizar (más) el otro (Holmström y Milgrom, 1994; Roberts, 2004).³⁰ Por lo tanto, si se introduce un instrumento, se vuelve más atractivo introducir (o incrementar el uso de) un instrumento complementario. Para ilustrar la idea de complementariedad, consideremos el caso de Lincoln Electric (Fast y Berg, 1975; Milgrom y Roberts, 1995), un fabricante de equipos para soldadura con un amplio dominio de la industria en Estados Unidos.³¹ Lincoln Electric tiene un récord de performance difícil de igualar: dio ganancias en cada trimestre de cada año durante más de un siglo y en Estados Unidos nunca ha dado pérdidas ni despedido a un solo trabajador. Gigantes como General Electric y Westinghouse eligieron abandonar la industria antes que intentar competir con Lincoln. La piedra angular de la estrategia de Lincoln es el uso extendido del pago a destajo o por unidad. Como vimos en las secciones 2 y 3, el pago por unidad provee incentivos poderosos para el trabajo duro, pero tiene sus desventajas —allí es donde las otras prácticas (complementarias) de Lincoln Electric cobran sentido.

Por ejemplo, en un proceso de producción en cadena o en equipo, la capacidad de los trabajadores de responder a los incentivos es limitada; en Lincoln, el sistema de producción está diseñado para que cada trabajador pueda manejar sus tiempos —a costa de tolerar grandes inventarios de piezas no terminadas. Los incentivos fuertes a la cantidad suelen entrar en conflicto con la calidad; Lincoln hace que cada trabajador grabe su nombre en la máquina sobre la que trabajó y, si la máquina falla, el trabajador debe repararla en su propio tiempo. Y siempre está el problema de vender todo lo que el

³⁰ En términos matemáticos, dos variables son complementarias cuando el rendimiento marginal de una es creciente en el nivel de la otra (es decir, la derivada cruzada es positiva). Para el caso general, no diferenciable, consultar la contribución original de Holmström y Milgrom (1994).

³¹ La descripción que sigue está basada en Roberts (2004).

sistema incentiva a producir en momentos en que las ventas están deprimidas. Para lidiar con esto, la empresa limita el tiempo que los trabajadores pueden pasar en sus puestos e impone horas extras cuando la demanda lo justifica.

Otras actividades valiosas (como ayudar a otros o aportar ideas novedosas), pero difíciles de cuantificar, son desalentadas por el sistema de pago a destajo. Para hacer este sistema más efectivo, Lincoln lo complementa con un esquema de bonus individuales. El monto del bonus es determinado subjetivamente por el supervisor del empleado sobre la base de la calidad de lo producido y el nivel de colaboración mostrado, entre otros factores, y puede duplicar la compensación estándar del trabajador. Un problema adicional con el pago por unidad aparece cuando los trabajadores desconfían del management y temen que el pago por unidad sea reducido si revelan qué tan productivos pueden ser. Para evitar este problema, la compañía promete que el pago por unidad no será modificado a menos que haya un cambio tecnológico y permite que el trabajador apele cualquier cambio.³² Para hacer la promesa creíble, Lincoln ha desarrollado toda una cultura corporativa que alienta la confianza de los trabajadores hacia el management: la empresa es, esencialmente, de propiedad de los empleados, y siempre ha mantenido una política de comunicación abierta entre niveles jerárquicos; además, los managers suelen ser reclutados dentro de la propia compañía, no tienen estacionamiento privilegiado en las plantas, su compensación no es tanto mayor que la de los trabajadores, almuerzan en el mismo lugar que los obreros y sus oficinas son de lo más sobrias –como para evitar un clima de “nosotros” contra “ellos” entre trabajadores y management.

³² Y para evitar temores de despidos, Lincoln promete no echar a nadie – y, ciertamente, ha desarrollado una reputación de cumplir esta promesa: nunca echó a nadie en más de 100 años. Para una simpática ilustración de los temores de los empleados asociados al pago por unidad, ver Roy (1952).

Las prácticas de Lincoln ilustran acabadamente la noción de sistema de incentivos: cada una de ellas por separado (como la promesa de no echar a nadie o la tolerancia hacia altos niveles de inventario de productos sin terminar) no parece tener mayor sentido, pero tomadas en conjunto rinden mucho más de lo esperado –precisamente, porque la complementariedad entre los instrumentos hace que el todo sea más que la suma de las partes.³³ La complementariedad entre prácticas implica un patrón coherente, en el sentido de que ningún cambio, por grande que sea, en sólo una parte de los instrumentos puede mejorar la performance de la organización (Roberts, 2004). Ello podría explicar por qué el modelo de Lincoln es único y nadie ha podido copiarlo, a pesar de ser un caso tan conocido y estudiado (el estudio de caso de Fast y Berg, 1975, es el más vendido de la historia de Harvard Business School): no alcanza con imitar algunas de las prácticas, sino que el sistema completo debe ser reproducido para arribar a un patrón coherente.^{34,35}

El rol del manager II

La idea de sistema tiene aplicabilidad más allá de los contratos de incentivos. De hecho, aplicando la misma lógica, el diseño global de los incentivos de una organización

³³ Para evidencia más sistemática en este sentido, el lector interesado puede consultar, por ejemplo, Ichniowski et al. (1997) y Bresnahan et al. (2002).

³⁴ En los '80, General Motors intentó copiar el modelo industrial de Toyota (por ejemplo, introduciendo sistemas de producción automatizados) pero retuvo algunas de sus prácticas usuales (como una variedad reducida de productos). El resultado fue un patrón incoherente que redundó en su momento en la mayor pérdida anual para una corporación en la historia de Estados Unidos.

³⁵ Aparentemente, ni la propia Lincoln Electric es capaz de copiarse a sí misma: ver Hastings (1999) para una narración de los problemas enfrentados por la compañía cuando intentó expandir sus operaciones fuera de Estados Unidos.

debería complementar sus prácticas de recursos humanos, producción, desarrollo de productos y competencia (Gibbons, 1998). Esta es la responsabilidad fundamental de los gerentes generales. No sólo deben concebir una estrategia para la organización, sino también crear una organización que implemente esa estrategia en el contexto dado en el que la organización opera. En este capítulo hemos analizado sólo algunas de las lecciones de la Economía de las Organizaciones respecto del problema del diseño organizacional. El campo es mucho más amplio y el lector interesado en profundizar en estos temas debería comenzar con el libro de Roberts (2004) mientras aguarda la edición del *Handbook of Organizational Economics* de Gibbons y Roberts (2011). En el ínterin, espero que este capítulo le haya dado algo en que pensar (y ganas de seguir leyendo)...

6. Referencias bibliográficas

- Arrow, Kenneth J. (1974). *The Limits of Organization*, New York: W. W. Norton & Company.
- Arrow, Kenneth J. y Frank H. Hahn (1971). *General Competitive Analysis*, San Francisco: Holden Day.
- Baker, George P. (1992). "Incentive Contracts and Performance Measurement", *The Journal of Political Economy* 100(3): 598-614.
- Baker, George P. (2002). "Distortion and Risk in Optimal Incentive Contracts", *The Journal of Human Resources* 37(4): 728-751.
- Baker, George P., Robert Gibbons y Kevin J. Murphy (1994). "Subjective Performance Measures in Optimal Incentive Contracts", *The Quarterly Journal of Economics* 109(4): 1125-1156.

- Baker, George P., Robert Gibbons y Kevin J. Murphy (2002). “Relational Contracts and the Theory of the Firm”, *The Quarterly Journal of Economics* 117(1): 39-84.
- Baker, George P., Michael C. Jensen y Kevin J. Murphy (1988). “Compensation and Incentives: Practice vs. Theory”, *The Journal of Finance* 43(3): 593-616.
- Bose, Arup, Debashis Pal y David E. M. Sappington (2011). “On the Performance of Linear Contracts”, *Journal of Economics & Management Strategy* 20(1): 159--193.
- Bresnahan, Timothy F., Erik Brynjolfsson y Lorin M. Hitt (2002). “Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence”, *The Quarterly Journal of Economics* 117(1): 339-376.
- Clausing, Kimberly A. (2003). “Tax-Motivated Transfer Pricing and US Intrafirm Trade Prices”, *Journal of Public Economics* 87 (9-10): 2207-2223.
- Coase, Ronald H. (1937). “The Nature of the Firm”, *Economica* 4(16): 386-405.
- Coase, Ronald H. (1960). “The Problem of Social Cost”, *Journal of Law and Economics* 3: 1-44.
- Courty, Pascal y Gerald Marschke (2004). “An Empirical Investigation of Gaming Responses to Explicit Performance Incentives”, *Journal of Labor Economics* 22(1): 23-56.
- Datar, Srikant, Susan Cohen Kulp y Richard A. Lambert (2001). “Balancing Performance Measures”, *Journal of Accounting Research* 39(1): 75-92.
- Fast, Norman A. y Norman D. Berg (1975). “The Lincoln Electric Company”, Harvard Business School Case #376-028.

- Feltham, Gerald A. y Jim Xie (1994). "Performance Measure Congruity and Diversity in Multi-Task Principal/Agent Relations", *The Accounting Review* 69(3): 429-453.
- Freund, Caroline y Diane Weinhold (2002). "The Internet and International Trade in Services", *The American Economic Review* 92(2): 236-240.
- Gibbons, Robert (1998). "Incentives in Organizations", *The Journal of Economic Perspectives* 12(4): 115-132.
- Gibbons, Robert (1999). "Taking Coase Seriously", *Administrative Science Quarterly* 44(1): 145-157.
- Gibbons, Robert (2003). "Team Theory, Garbage Cans and Real Organizations: Some History and Prospects of Economic Research on Decision-Making in Organizations", *Industrial and Corporate Change* 12(4): 753-787.
- Gibbons, Robert (2005a). "Four Formal(izable) Theories of the Firm?", *Journal of Economic Behavior & Organization* 58(2): 200--245.
- Gibbons, Robert (2005b). "Incentives Between Firms (and Within)", *Management Science* 51(1): 2-17.
- Gibbons, Robert (2006). "What the Folk Theorem Doesn't Tell Us", *Industrial and Corporate Change* 15(2): 381-386.
- Gibbons, Robert (2010a). "Inside Organizations: Pricing, Politics, and Path Dependence", *Annual Review of Economics* 2: 337-365.
- Gibbons, Robert (2010b). Lecture Notes for the Organizational Economics and Corporate Strategy MBA Course (15.903), MIT Sloan School of Management.
- Gibbons, Robert y John Roberts (2011). *The Handbook of Organizational Economics*, Princeton: Princeton University Press.

- Granovetter, Mark (1985). "Economic Action and Social Structure: The Problem of Embeddedness", *The American Journal of Sociology* 91(3): 481-510.
- Grossman, Sanford J. y Oliver D. Hart (1983). "An Analysis of the Principal-Agent Problem", *Econometrica* 51(1): 7-45.
- Hart, Oliver D. (2008). "Reference Points and the Theory of the Firm", *Economica* 75(299): 404-411.
- Hart, Oliver D. y Bengt Holmström (1987). "The Theory of Contracts", en T. Bewley (ed.), *Advances in Economic Theory: Fifth World Congress*, Cambridge: Cambridge University Press.
- Hastings, Donald F. (1999). "Lincoln Electric's Harsh Lessons from International Expansion", *Harvard Business Review* (mayo-junio): 163-178.
- Holmström, Bengt (1979). "Moral Hazard and Observability", *The Bell Journal of Economics* 10(1): 74-91.
- Holmström, Bengt y Paul Milgrom (1987). "Aggregation and Linearity in the Provision of Intertemporal Incentives", *Econometrica* 55(2): 303-328.
- Holmström, Bengt y Paul Milgrom (1991). "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design", *Journal of Law, Economics, & Organization* 7(Special Issue): 24-52.
- Holmström, Bengt y Paul Milgrom (1994). "The Firm as an Incentive System", *The American Economic Review* 84(4): 972-991.
- Ichniowski, Casey, Kathryn Shaw y Giovanna Prennushi (1997). "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines", *The American Economic Review* 87(3): 291-313.

- Kerr, Steven (1975). "On the Folly of Rewarding A, While Hoping for B", *Academy of Management Journal* 18(4): 769-783.
- Laffont Jean-Jacques y David Martimort (2002). *The Theory of Incentives. The Principal-Agent Model*, Princeton: Princeton University Press.
- Lazear, Edward P. (2000). "Performance Pay and Productivity", *The American Economic Review* 90(5): 1346-1361.
- Levin, Jonathan (2003), "Relational Incentive Contracts", *The American Economic Review* 93(3): 835-857.
- Macaulay, Stewart (1963). "Non-Contractual Relations in Business: A Preliminary Study", *American Sociological Review* 28(1): 55-67.
- MacDonald, Glenn M. (1984). "New Directions in the Economic Theory of Agency", *The Canadian Journal of Economics* 17(3): 415-440.
- McMillan, John (2002). *Reinventing the Bazaar: The Natural History of Markets*, New York: W. W. Norton & Company.
- Milgrom, Paul y John Roberts (1995). "Complementarities and Fit. Strategy, Structure, and Organizational Change in Manufacturing", *Journal of Accounting & Economics* 19(2-3): 179-208.
- Oyer, Paul (1998). "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect of Business Seasonality", *The Quarterly Journal of Economics* 113(1): 149-185.
- Pigou, Arthur C. (1922). "Empty Economic Boxes: A Reply", *The Economic Journal* 32(128): 458-465.
- Pigou, Arthur C. (1932). *The Economics of Welfare*, London: Macmillan.

Prendergast, Canice (2000). “What Trade-Off of Risk and Incentives?”, *The American Economic Review* 90(2): 421-425.

Roberts, John (2004). *The Modern Firm*, Oxford: Oxford University Press.

Roy, Donald (1952). “Quota Restriction and Goldbricking in a Machine Shop”, *The American Journal of Sociology* 57(5): 427-442.

Simon, Herbert A. (1991). “Organizations and Markets”, *The Journal of Economic Perspectives* 5(2): 25-44.

Williamson, Oliver (1985). *The Economic Institutions of Capitalism*, New York: Free Press.

Intermediarios de certificación y transmisión de información en mercados con información asimétrica

*Lucía Quesada**
(Departamento de Economía
Universidad Torcuato Di Tella)

1. Introducción

Las asimetrías de información tienen importantes consecuencias en la asignación de recursos de una economía. En muchos casos, los agentes informados se ven obligados a llevar a cabo acciones costosas para enviar al resto de los agentes una señal sobre su información. Ejemplos de estas acciones costosas son las garantías de calidad o la inversión en establecer una buena reputación. Sin embargo, en muchos casos estas acciones ofrecen una solución sólo parcial al problema de la información asimétrica. En este contexto, surge una nueva institución de mercado, los intermediarios de certificación, cuyo principal objetivo consiste en reducir las asimetrías de información. El rol de estos intermediarios es recabar información de los agentes informados y transmi-

* Universidad Torcuato Di Tella, Departamento de economía, Sáenz Valiente 1010, C1428BIJ Buenos Aires, Argentina. Email: *lquesada@utdt.edu*.

tirla, al menos en parte, al resto de los agentes. Por lo tanto, los intermediarios de certificación prestan un servicio potencialmente importante, si ayuda a eliminar ineficiencias en la asignación de recursos.

Podemos pensar en varios ejemplos de instituciones que cumplen este rol en mercados de distinto tipo. Entre otros, los laboratorios de prueba para los bienes industriales, los servicios de auditoría, las agencias calificadoras de riesgo, las concesionarias de autos usados, los bancos de inversión y las instituciones educativas certifican la calidad de los productos (bienes industriales, firmas que buscan inversores, autos usados, alumnos).

En este sentido, es importante entender el comportamiento de los intermediarios de certificación en lo que se refiere a sus políticas de transmisión de información y de tarificación. En este capítulo vamos a hacer una pequeña selección de modelos que intentan abordar distintos aspectos de la problemática de los intermediarios de certificación. Para eso, vamos a plantear un mercado en el que conviven tres tipos de individuos. Primero, el vendedor de un producto de una determinada calidad, que es su información privada. Segundo, potenciales compradores de ese producto que sólo tienen información difusa sobre la calidad (sólo conocen la distribución de probabilidad). Por último, un intermediario de certificación con acceso a una tecnología que le permite recabar información sobre el producto y una reputación que le permite transmitir información de manera creíble.

Con esta estructura, vamos a presentar tres aspectos distintos del problema de la transmisión de información por parte de los intermediarios de certificación. En el primer modelo, analizamos cómo puede beneficiarse el intermediario con la manipulación de la información. En el segundo modelo, vamos a mostrar que la manipulación de la información puede no ser una política óptima si el intermediario

puede renegociar el contrato por el servicio de certificación. En particular, vamos a mostrar que el intermediario puede producir “demasiada” información, en el sentido de que la disposición a pagar por esa información es menor al costo de obtenerla. Finalmente, en el tercer modelo vamos a investigar los incentivos que puede tener el intermediario para coludir con el vendedor del producto, con el objetivo de transmitir información falsa que permita vender el producto a un precio más alto.

Si bien el radio de acción de los intermediarios de certificación es muy amplio, la literatura microeconómica que estudia el comportamiento de estos agentes es relativamente reciente. Existe una literatura anterior, que investiga el rol de intermediarios en general, como Rubinstein y Wolinsky (1987) y Biglaiser (1993). Sin embargo, el primer artículo de referencia en el tema de los intermediarios de certificación es Lizzeri (1999), que desarrolla un modelo en el que un intermediario monopólico decide sobre el precio del servicio de certificación y la política de transmisión de información. En la Sección 3 de este trabajo, presentamos una versión simplificada de ese modelo. El resultado más importante de Lizzeri (1999) es que en el único equilibrio el intermediario se apropia de todo el excedente del mercado, sin transmitir ningún tipo de información a los potenciales compradores. Es decir, el servicio de certificación funciona exactamente como una acción costosa de señalización. Los compradores interpretan la ausencia de certificación como un indicador de la peor calidad posible. Albano y Lizzeri (2001) introducen un problema de riesgo moral en el cual la calidad del producto es el resultado de un esfuerzo de inversión por parte del vendedor. Muestran que la certificación puede dar los incentivos correctos para invertir en calidad. Faure-Grimaud, Peyrache y Quesada (2009) muestran que si el intermediario no puede comprometerse a no renegociar el

contrato de certificación, cuando la calidad es alta la información siempre se revela. Más aún, bajo determinadas condiciones, toda la información se revela en equilibrio. La introducción de competencia entre intermediarios reduce la cantidad de información que se revela.

Gu (2008) analiza los efectos de incorporar una tecnología de testeo imperfecta para el intermediario. Muestra que, incluso en este caso, existen equilibrios en los que sólo los vendedores de productos de alta calidad eligen contratar al intermediario.

Strausz (2005) y Peyrache y Quesada (2011) analizan el problema de la potencial colusión entre el intermediario y el vendedor. La diferencia fundamental entre ambos artículos es que el primero se concentra sólo en los casos en los que existe un equilibrio sin colusión. Peyrache y Quesada (2011) muestran que si el valor de la reputación del intermediario es relativamente bajo, existen equilibrios en los que el intermediario colude cuando la calidad es baja. Esto significa que revela información que hace creer a los compradores que la calidad del producto es mayor a la real.

El capítulo está organizado de la siguiente manera. En la Sección 2 describimos las características generales del mercado que vamos a analizar. En la Sección 3 presentamos una versión simplificada de Lizzeri (1999) y mostramos que en el único equilibrio posible el intermediario no revela ningún tipo de información sobre la calidad del producto. Sólo revela si el vendedor adquirió el servicio de certificación o no. Es decir, el intermediario sólo funciona como una señal de la calidad del producto. El mercado considera que si el vendedor no contrata al intermediario es porque su calidad es muy baja. En particular, si la revelación de información no contribuye a aumentar la eficiencia en el intercambio, el vendedor siempre decidirá contratar al intermediario. En cambio, si la revelación de información aumenta la eficien-

cia, el vendedor contratará al intermediario sólo si la calidad del producto es tal que la disposición a pagar de los compradores es mayor al costo de producción. En la Sección 4 presentamos una variante de Faure-Grimaud, Peyrache y Quesada (2009) y mostramos que, cuando el intermediario no tiene capacidad para comprometerse a no renegociar el contrato de certificación, en cualquier equilibrio la información buena siempre se revela. La razón es la siguiente. Supongamos que el intermediario se había comprometido a no revelar la información. Una vez realizado el test, descubre que la calidad del producto es muy alta. Esto quiere decir que la disposición a pagar de los compradores aumentaría si decidiera revelar la información, porque impide que este producto sea confundido con otros de menor calidad. Entonces, está tentado de ofrecer una renegociación del contrato al vendedor en la que le cobra un precio extra por revelar la información, para captar parte de ese aumento en la disposición a pagar. Es más, bajo determinadas condiciones (intermediario monopolístico y distribución uniforme de calidades), toda la información se revela en equilibrio. Si se introduce competencia en el mercado de intermediación, la cantidad de información que se revela en equilibrio disminuye. En la Sección 5 discutimos el problema de la colusión entre el intermediario y el vendedor, usando el modelo de Peyrache y Quesada (2011). Para eso introducimos un intermediario que se interesa por sus beneficios futuros (reputación). Mostramos que si el valor de la reputación es alto, existe un equilibrio en el que el intermediario nunca colude con el vendedor. Por el contrario, si el valor de la reputación es bajo, el intermediario siempre decide coludir cuando la calidad del producto es suficientemente baja. Además, mostramos que el precio de la certificación es una función creciente del valor de la reputación. Esto implica que si el intermediario valora mucho la reputación, sólo los productos

de calidad más alta se harán certificar, lo que reduce los incentivos a coludir. Finalmente, en la Sección 6 mencionamos otros aspectos del mercado de intermediarios de certificación que por cuestiones de espacio no fueron desarrollados en este trabajo.

2. El modelo

Consideramos un mercado con un vendedor de una unidad de un producto y $n \geq 2$ compradores.

Sea $\theta \in [a, b]$ la calidad del producto. Es conocida por el vendedor y desconocida por los compradores. A lo largo de este capítulo llamaremos θ al tipo del vendedor. Los compradores saben que la función de distribución acumulada de θ es F , con soporte completo en $[a, b]$. Suponemos que $b > a \geq 0$.

La utilidad de un comprador que compra un producto de calidad θ al precio β es $u_b = \theta - \beta$. Si no compra, su utilidad es 0. El producto se subasta entre los compradores. Como la información es simétrica entre los compradores y la valoración del bien es la misma para todos, el precio de venta de equilibrio es igual al valor esperado de θ dada la información disponible.

El costo de producción es 0 y la utilidad del vendedor es $u_s = \beta$ cuando vende el producto al precio β y 0 cuando no lo vende. En particular, la utilidad del vendedor es independiente del valor de θ , lo que elimina problemas de tipo *lemon*.¹ Esto implica que sin la existencia del intermediario en este mercado, el producto se vende con probabilidad 1 al

¹ Ver Akerlof (1970). Cuando el valor para el vendedor también aumenta con la calidad, las asimetrías de información pueden llevar a la desaparición del mercado.

precio $\beta = E(\theta) > 0$ y el vendedor se apropia de todo el excedente del mercado.

El intermediario, posee una tecnología que le permite testear el producto y conocer perfectamente el valor de θ . El costo del test es 0. El intermediario ofrece un contrato de certificación al vendedor, con un precio p por el servicio de certificación. La utilidad del intermediario es $\pi = p$ si el vendedor acepta y 0 en caso contrario.

Suponemos que el resultado del test genera evidencia que puede mostrarse para avalar la información revelada. A este supuesto se lo conoce en la literatura como información sólida (*hard information*). La consecuencia importante de este supuesto es que el intermediario puede ocultar el resultado del test pero no puede falsificarlo. Es decir, si revela alguna información, esa información tiene que ser verdadera.

Desarrollo temporal del juego

El intermediario ofrece un contrato de certificación $CC = (p, \alpha)$ al vendedor. El contrato establece un precio a pagar por el servicio de certificación, p y una probabilidad de hacer público el resultado del test α .²

La naturaleza elige el valor de θ y éste es revelado al vendedor.

El vendedor acepta (A) o rechaza (R) el contrato ofrecido por el intermediario.

Si el vendedor aceptó el contrato, el intermediario realiza el test y descubre θ .

² El contrato que propone Lizzeri (1999) contempla una variedad mucho mayor de reglas de revelación de información, que pueden incluso depender del resultado del test. Los resultados son los mismos que con esta regla más sencilla.

El intermediario revela la información de acuerdo con el contrato celebrado con el vendedor.

El producto se subasta entre los potenciales compradores.

3. El equilibrio sin revelación de información

En esta sección vamos a introducir el modelo más sencillo para empezar a entender los efectos que tiene la introducción del intermediario en el mercado, en cuanto a revelación de información y distribución del excedente total. Para eso, vamos a empezar suponiendo que el intermediario tiene la capacidad de comprometerse a ejecutar cualquier contrato (p, α) , aun cuando la política de revelación de información resulte subóptima una vez realizado el test.

Este es un juego secuencial con información incompleta y, por lo tanto, el concepto de equilibrio relevante es el de equilibrio bayesiano perfecto. Recordemos que en este tipo de equilibrio, las creencias de los jugadores deben actualizarse de acuerdo con la regla de Bayes si las acciones observadas son estrategias de equilibrio, pero son arbitrarias para acciones que no deberían jugarse en equilibrio.

En este mercado, las creencias de los compradores son importantes porque determinan el precio que se pagará por el producto al momento del intercambio. Dado que la utilidad del vendedor es creciente en β , su disposición a pagar por el servicio de certificación será mayor cuanto más favorables sean las creencias de los compradores. Es importante destacar que existen dos tipos de información que pueden modificar las creencias de los compradores. En primer lugar, la decisión del vendedor de contratar o no el servicio de certificación. En efecto, en equilibrio un subconjunto de tipos aceptará el contrato de certificación y el subconjunto complementario lo rechazará. Por lo tanto la decisión del vendedor permite a los compradores ubicar al vendedor en un

subconjunto o en el otro. En segundo lugar, la información que revele el intermediario influye en las creencias de los compradores.

El principal resultado de este modelo es que existe un único equilibrio en el que el intermediario se lleva todo el excedente apropiable del mercado pero en ningún caso revela la información sobre θ .

Supongamos que el intermediario ofrece cualquier contrato (p, α) . Con este contrato, los compradores tienen potencialmente dos fuentes de información. Primero, la eventual publicación del resultado del test, que les permitiría identificar con certeza el valor de θ . En ese caso, los compradores están dispuestos a pagar hasta θ por el producto. Segundo, si el resultado del test no es publicado, la decisión del vendedor de aceptar o rechazar el contrato. El máximo precio que los compradores están dispuestos a pagar en este caso, será $E(\theta | A)$ si el vendedor aceptó y $E(\theta | R)$ si el vendedor rechazó.

La utilidad esperada de un vendedor de tipo θ que acepta el contrato es, entonces

$$u_s(\theta, A) = \alpha\theta + (1 - \alpha)E(\theta | A) - p.$$

La utilidad esperada si rechaza el contrato es

$$u_s(\theta, R) = E(\theta | R).$$

Esto implica que el vendedor aceptará el contrato siempre que

$$\alpha\theta + (1 - \alpha)E(\theta | A) - p \geq E(\theta | R),$$

o

$$\alpha(\theta - E(\theta|A)) + E(\theta|A) - E(\theta|R) \geq p.$$

Como el lado derecho es creciente en θ , existirá un umbral $\tilde{\theta}$ tal que el vendedor acepta el contrato si y sólo si $\theta \geq \tilde{\theta}$. Usando la regla de Bayes, tenemos entonces que en equilibrio debe darse que

$$E(\theta|A) = E(\theta|\theta \geq \tilde{\theta}),$$

$$E(\theta|R) = E(\theta|\theta < \tilde{\theta}).$$

Es más, el umbral $\tilde{\theta}$ es el que está indiferente entre aceptar y rechazar y está definido por la igualdad

$$\alpha(\tilde{\theta} - E(\theta|\theta \geq \tilde{\theta})) + E(\theta|\theta \geq \tilde{\theta}) - E(\theta|\theta < \tilde{\theta}) = p.$$

Veamos ahora la decisión del intermediario. El beneficio del intermediario con un contrato (p, α) es

$$\pi(p, \alpha) = \int_{\tilde{\theta}}^b p dF(\theta) = p[1 - F(\tilde{\theta})],$$

con

$$p = \alpha(\tilde{\theta} - E(\theta|\theta \geq \tilde{\theta})) + E(\theta|\theta \geq \tilde{\theta}) - E(\theta|\theta < \tilde{\theta}).$$

Podemos escribir el problema del intermediario como

$$\max_{\tilde{\theta}, \alpha} \int_{\tilde{\theta}}^b [\alpha(\tilde{\theta} - E(\theta|\theta \geq \tilde{\theta})) + E(\theta|\theta \geq \tilde{\theta}) - E(\theta|\theta < \tilde{\theta})] dF(\theta).$$

Esta función es decreciente en α porque $\tilde{\theta} < E(\theta | \theta \geq \tilde{\theta})$. Por lo tanto, es óptimo elegir $\alpha = 0$. Dado este resultado, la elección del intermediario se reduce a

$$\max_{\tilde{\theta}} \int_{\tilde{\theta}}^b [E(\theta | \theta \geq \tilde{\theta}) - E(\theta | \theta < \tilde{\theta})] dF(\theta),$$

que es equivalente a

$$\max_{\tilde{\theta}} [1 - F(\tilde{\theta})][E(\theta | \theta \geq \tilde{\theta}) - E(\theta | \theta < \tilde{\theta})].$$

Finalmente, sabemos que $\forall \tilde{\theta}$,

$$[1 - F(\tilde{\theta})]E(\theta | \theta \geq \tilde{\theta}) + F(\tilde{\theta})E(\theta | \theta < \tilde{\theta}) = E(\theta).$$

Por lo tanto, el problema puede escribirse como

$$\max_{\tilde{\theta}} [E(\theta) - E(\theta | \theta < \tilde{\theta})].$$

Esta función es decreciente en $\tilde{\theta}$ y por lo tanto es óptimo elegir $\tilde{\theta} = a$. Con este resultado y $\alpha = 0$ obtenemos que

$$p = E(\theta) - a,$$

donde $a = E(\theta | R)$ es la creencia sobre el valor de θ si el vendedor rechaza el contrato. En realidad, con $\tilde{\theta} = a$ estas creencias son creencias fuera del equilibrio, pero debe ser que $E(\theta | R) = a$ para garantizar que las estrategias formen un equilibrio. En efecto, supongamos que $E(\theta | R) > a$. En este caso, $\pi = p = E(\theta) - E(\theta | R)$. Entonces, $\exists \varepsilon > 0$, peque-

ño tal que $\alpha' = \varepsilon$, $\tilde{\theta}' = a + \varepsilon$ y $p' = \varepsilon(a + \varepsilon) + (1 - \varepsilon)\mathbf{E}(\theta | \theta \geq a + \varepsilon) - \mathbf{E}(\theta | \theta < a + \varepsilon)$. Para $\varepsilon \rightarrow 0$, el beneficio es $\pi' \rightarrow \mathbf{E}(\theta) - a > \pi$.

Proposición 1 *Existe un único equilibrio del juego de certificación con las siguientes características:*

- *El intermediario ofrece un contrato $(p, \alpha) = (\mathbf{E}(\theta) - a, 0)$.*
- *El vendedor acepta el contrato cualquiera sea su tipo.*
- *Los compradores pagan un precio $\beta(A) = \mathbf{E}(\theta)$ si el vendedor aceptó el contrato y $\beta(R) = a$ si el vendedor rechazó el contrato.*
- *Las creencias de los compradores vienen dadas por $\theta \sim F[a, b]$ si el vendedor acepta el contrato y $\theta = a$ con probabilidad 1 si el vendedor rechaza el contrato.*

Lo curioso de este contrato es que el intermediario termina extrayendo todo el excedente posible sin prestar en definitiva ningún servicio, porque no revela ningún tipo de información. La razón por la que puede extraer todo el excedente es que si el vendedor decide no contratar al intermediario, el mercado interpreta esa decisión como una señal de que el vendedor es el peor tipo posible. De hecho, éstas son las únicas creencias fuera del equilibrio que sostienen este equilibrio.

En el único equilibrio de este modelo, la única información que se revela es que el vendedor compró el servicio de certificación. La existencia del intermediario no reduce las asimetrías de información. Simplemente, traslada excedente del vendedor al intermediario. El vendedor acepta pagar el

precio de la certificación porque si no lo hiciera sería considerado por los compradores como el peor de los tipos posibles. El intermediario prefiere no revelar información porque si revelara información se vería obligado a cobrar un precio más bajo al tipo umbral. Podría pensarse que eso está asociado con la particular regla de revelación de información que supone este modelo (por ejemplo, que el intermediario revele o no la información no depende del resultado del test). Sin embargo, Lizzeri (1999) muestra que este resultado se generaliza para cualquier regla de revelación del tipo $D : [\sigma, b] \rightarrow Q$, siendo Q el conjunto de distribuciones de probabilidad de Borel sobre números reales. La única condición que se necesita para generalizar el resultado es que la distribución de θ sea tal que $E(\theta | \theta \geq x) - E(\theta | \theta \leq x) \geq E(\theta) - \sigma, \forall x$.³

La clave para explicar este resultado es que la información en este modelo no tiene valor social. El excedente total promedio en este mercado es $E(\theta)$ tanto con información completa como con información asimétrica. Esto es porque $\sigma \geq 0$ y entonces el intercambio es siempre eficiente (para cualquier valor posible de θ el valor para los compradores es mayor que para el vendedor). La existencia del intermediario sólo afecta la distribución del excedente total, pero no su magnitud.

Información con valor social: estándares mínimos

Para analizar un caso en el que la revelación de información aumenta el excedente total, supongamos que $\sigma < 0$.⁴

³ Muchas distribuciones satisfacen esta propiedad, entre ellas, la distribución uniforme. También, cualquier distribución con densidad decreciente satisface esta propiedad.

⁴ Una manera alternativa y equivalente en cuanto a los resultados, sería suponer que existe un costo de producción superior a σ .

Esto significa que con información asimétrica el intercambio es ineficiente con probabilidad positiva. En particular, el intercambio se realiza si y sólo si $E(\theta) \geq 0$, independientemente de que la realización de θ sea positiva o negativa.

Supongamos que el intermediario ofrece algún contrato (p, α) . Si el intermediario revela el resultado del test, el intercambio ocurrirá sólo si $\theta > 0$. En ese caso, los compradores pagarán un precio igual a θ . Si $\theta \leq 0$, el intercambio no se realiza. Entonces, la utilidad de un vendedor tipo θ que acepta el contrato es

$$u_s(\theta, A) = \alpha \max\{\theta, 0\} + (1 - \alpha) \max\{E(\theta | A), 0\} - p.$$

Su utilidad si rechaza el contrato es

$$u_s(\theta, R) = \max\{E(\theta | R), 0\}.$$

Por lo tanto, aceptará el contrato si

$$p \leq \max\{\alpha\theta + (1 - \alpha)E(\theta | A), 0\} - \max\{E(\theta | R), 0\}.$$

Nuevamente, esta condición es creciente en θ y existe un umbral $\tilde{\theta}$ tal que el vendedor acepta si y sólo si $\theta \geq \tilde{\theta}$, definido implícitamente por la condición

$$p = \max\{\alpha\tilde{\theta} + (1 - \alpha)E(\theta | \theta \geq \tilde{\theta}), 0\} - \max\{E(\theta | \theta < \tilde{\theta}), 0\}.$$

Ahora podemos mostrar que en equilibrio $\tilde{\theta} : \alpha\tilde{\theta} + (1 - \alpha)E(\theta | \theta \geq \tilde{\theta}) > 0$. Si no fuera así, el precio de la certificación y los beneficios del intermediario serían nulos. Entonces, el intermediario podría subir un poco el precio, aumentando el umbral y aumentar sus beneficios.

El beneficio del intermediario es

$$\begin{aligned}\pi &= \int_{\tilde{\theta}}^b [\alpha \tilde{\theta} + (1-\alpha)E(\theta | \theta \geq \tilde{\theta}) - \max\{E(\theta | \theta < \tilde{\theta}), 0\}] dF(\theta) \\ &= (1-F(\tilde{\theta}))[\alpha \tilde{\theta} + (1-\alpha)E(\theta | \theta \geq \tilde{\theta}) - \max\{E(\theta | \theta < \tilde{\theta}), 0\}],\end{aligned}$$

decreciente en α . Por lo tanto, es óptimo para el intermediario elegir $\alpha = 0$. Reescribimos entonces el beneficio como

$$\pi = (1 - F(\tilde{\theta})) [E(\theta | \theta \geq \tilde{\theta}) - \max\{E(\theta | \theta < \tilde{\theta}), 0\}].$$

Queda por elegir el valor óptimo de $\tilde{\theta}$. Para eso, definimos $\underline{\theta} > 0$: $E(\theta | \theta < \underline{\theta}) = 0$. Quedan dos casos por analizar.

Caso a: $\tilde{\theta} \geq \underline{\theta}$. En este caso tenemos que $E(\theta | \theta < \tilde{\theta}) \geq 0$ y el beneficio del intermediario es

$$\begin{aligned}\pi &= (1 - F(\tilde{\theta})) [E(\theta | \theta \geq \tilde{\theta}) - E(\theta | \theta < \tilde{\theta})] \\ &= E(\theta) - E(\theta | \theta < \tilde{\theta}),\end{aligned}$$

decreciente en $\tilde{\theta}$. Por lo tanto, es óptimo elegir $\tilde{\theta} = \underline{\theta}$ y el beneficio del intermediario es

$$\begin{aligned}\pi_a &= (1 - F(\underline{\theta})) E(\theta | \theta \geq \underline{\theta}) \\ &= \int_{\underline{\theta}}^b \theta dF(\theta).\end{aligned}$$

Caso b: $\tilde{\theta} < \underline{\theta}$. En este caso tenemos que $E(\theta | \theta < \tilde{\theta}) < 0$ y el beneficio del intermediario es

$$\begin{aligned}\pi &= (1 - F(\tilde{\theta}))E(\theta | \theta \geq \tilde{\theta}) \\ &= \int_{\tilde{\theta}}^b \theta dF(\theta),\end{aligned}$$

que tiene un máximo en $\tilde{\theta} = 0 < \underline{\theta}$. El beneficio del intermediario es

$$\begin{aligned}\pi_b &= (1 - F(0))E(\theta | \theta \geq 0) \\ &= \int_0^b \theta dF(\theta).\end{aligned}$$

Dado que $\underline{\theta} > 0$, resulta que $\pi_b > \pi_a$. Por lo tanto, es óptimo elegir

$$\tilde{\theta} = 0.$$

Esto implica que en el equilibrio

$$p = E(\theta | \theta \geq 0).$$

Esta discusión demuestra que en el único equilibrio del juego, el intermediario ofrece un contrato $(p, \alpha) = (E(\theta | \theta \geq 0), 0)$ y el vendedor acepta si y sólo si $\theta \geq 0$.

El aspecto interesante de este equilibrio es que si bien $\alpha = 0$ (no hay revelación directa de información), la decisión del vendedor revela si el intercambio es eficiente o no. Por lo tanto, con esta información los compradores sólo compran productos certificados, i.e., aquéllos para los cuales el intercambio es eficiente.

Proposición 2 En el equilibrio con intermediario, el intercambio es siempre eficiente.

Cuando $\alpha < 0$ la existencia del intermediario aumenta el excedente total, que pasa de $\max\{E(\theta), 0\}$ a $E(\theta | \theta \geq 0)$. Sin embargo, la transmisión de información de este intermediario es sólo indirecta, a través de la decisión del vendedor de comprar el servicio de certificación. En cualquier caso, la publicación de información directa por parte del intermediario es nula. Esto es equivalente a un servicio de certificación que establece un estándar mínimo, i.e., $\theta = 0$ y el resultado del test sólo arroja si se alcanzó o no ese estándar mínimo. En equilibrio, sólo aquéllos que saben que alcanzarán el estándar contratarán el servicio de certificación.

4. Posibilidad de renegociación e información imperfecta por parte del vendedor

En el modelo desarrollado anteriormente se realizaron dos supuestos poco atractivos. Primero, el intermediario es capaz de comprometerse a una política de revelación de información que puede ser ineficiente ex post. Sin esa capacidad de compromiso, el intermediario podría querer renegociar el contrato para extraer una renta adicional. Segundo, el vendedor conoce perfectamente el valor de θ antes de realizarse el test. Sin embargo, es razonable pensar que si bien el vendedor tiene cierta información, probablemente más precisa que los compradores, también aprende algo sobre θ con el resultado del test.

En esta sección vamos a modificar el modelo para poder incorporar estos dos aspectos y vamos a inspeccionar cuáles son los efectos sobre la política de revelación de información. Para simplificar, supongamos que θ está distribuida

uniformemente en el intervalo $[0,1]$.⁵ Supongamos también que el test tiene un costo $c \in [0, 1/2]$.⁶

En ese sentido, incorporamos tres nuevos supuestos al modelo.

1. Suponemos que el intermediario no puede comprometerse a no renegociar el contrato una vez realizado el test. Este supuesto reduce el conjunto de contratos posibles entre el intermediario y el vendedor a contratos que son robustos a la renegociación (*renegotiation-proof*).
2. El vendedor no conoce exactamente el valor de θ sino que sólo tiene una señal sobre ese valor, μ , que se distribuye uniformemente en el intervalo $[\theta - \sigma, \theta + \sigma]$, donde $\sigma \in [0, \frac{1}{2}]$ mide la precisión de la señal. Esto, en particular, significa que en promedio el vendedor conoce su tipo (el valor esperado de la señal es θ) y que $\forall \mu_2 > \mu_1$ y $\forall x \in [0, 1]$,

$$\Pr(\theta < x \mid \mu_1) > \Pr(\theta < x \mid \mu_2),$$

es decir un μ más alto es señal de un θ más alto (dominación estocástica de primer orden). El valor de σ , por otra parte, da una idea de la precisión de la información ex ante del vendedor. Cuanto menor sea σ , más precisa será la información del vendedor.

3. Los compradores no observan la decisión del vendedor de contratar al intermediario. Sólo pueden inferir que se contrató al intermediario cuando éste hace

⁵ Los resultados cualitativos no dependen de este supuesto, pero la exposición se simplifica notablemente.

⁶ Los resultados de la sección anterior no cambian si se incorpora un costo de realizar el test.

público el resultado del test. Este supuesto implica que cuando el intermediario no revela información, los compradores no saben si esto es porque efectivamente decidió no revelarla o porque el vendedor no se hizo certificar. Por lo tanto, las creencias de los compradores sólo dependen de la información revelada y no de la decisión del vendedor.

El primer supuesto introducido (imposibilidad de comprometerse a no renegociar), implica que el equilibrio en este modelo es necesariamente distinto al de la Sección 3, cualquiera sea σ . Para ilustrar esto, supongamos que $\sigma = 0$. En este caso, de acuerdo con la Sección 3, en el único equilibrio el intermediario ofrecería el contrato $(p, \alpha) = (1/2, 0)$ y el vendedor aceptaría el contrato, cualquiera sea su tipo. Supongamos ahora que $\theta > 1/2$. Cuando el intermediario realiza el test, descubre esta información y le ofrece al vendedor renegociar el contrato de la siguiente manera: $(p', \alpha') = (\tau, 1)$, con $\tau \in (0, \theta - 1/2)$. El intermediario gana con este nuevo contrato porque recibe un nuevo pago $\tau > 0$. El vendedor gana con esta renegociación, porque su verdadero tipo será revelado y podrá vender el producto a un precio de θ en lugar de venderlo a un precio de $1/2$ y el pago adicional $\tau < \theta - 1/2$. Por lo tanto, el equilibrio de Lizzeri (1999) no es robusto a la renegociación.

Un primer paso para entender cómo cambian las predicciones del modelo, es identificar las características que debe tener un contrato que es robusto a la renegociación. Vamos a suponer que el resultado del test que realiza el intermediario es observado simultáneamente por el intermediario y la firma.⁷ Esto implica que al momento de la potencial renego-

⁷ En teoría, uno podría diseñar un mecanismo compatible en incentivos ex post que indujera al intermediario a revelar el resultado del test al vendedor.

ciación, las dos partes tienen la misma información. Cuando la renegociación ocurre con información simétrica, las únicas asignaciones robustas a la renegociación son las asignaciones eficientes ex post.

Llamamos ϕ a la información que tienen los compradores cuando no se revela información. Recordemos que la no revelación de información transmite cierta información a los compradores: o bien el vendedor no aceptó el contrato, o la calidad del producto era tal que el contrato especificaba no revelar información. Dada esa información, los compradores asignan un valor a la calidad del producto cuando no se reveló información. Sea $\theta_\phi = E_\theta(\theta | \phi)$ ese valor, donde $E_\theta(x(\theta) | \phi) = \int_0^1 x(\theta) dF(\theta | \phi)$ es el valor esperado, cuando la esperanza se toma con respecto a θ . Con esto, podemos describir el conjunto de contratos robustos a la renegociación.

El precio que los compradores están dispuestos a pagar por el producto dada la regla de revelación de información es

$$\beta(\alpha) = \alpha\theta + (1 - \alpha)\theta_\phi.$$

La regla de revelación de información eficiente ex post es la que maximiza la disposición a pagar de los compradores. De este razonamiento, deducimos el siguiente resultado.

Proposición 3 Un contrato es robusto a la renegociación si y sólo si tiene la siguiente propiedad:

$$\alpha(\theta) = \begin{cases} 1 & \text{si } \theta \geq \theta_\phi, \\ 0 & \text{si } \theta < \theta_\phi. \end{cases}$$

Vemos, entonces, que las reglas de revelación de información compatibles con la posibilidad de renegociación

dependen del valor de θ . Por este motivo, para dar al intermediario la misma posibilidad de extraer todo el excedente que tenía en la Sección 3, vamos suponer que puede ofrecer contratos del tipo $\{p_0, p(\theta), \alpha(\theta)\}_\mu$ que estipulan una tarifa por adelantado p_0 , una tarifa que depende del resultado del test, $p(\theta)$, y la regla de revelación de información $\alpha(\theta)$. Además, permitimos que el contrato dependa (de una manera que sea compatible con los incentivos) de algún reporte que la firma haga de su señal μ .

Dado un contrato $\{p_0, p(\theta), \alpha(\theta)\}$ robusto a la renegociación, la utilidad del vendedor con señal μ que acepta el contrato es

$$E_\theta[(\alpha(\theta)\theta + (1 - \alpha(\theta))\theta_\phi - p(\theta) - p_0) | \mu].$$

Por otra parte, la utilidad del vendedor con señal μ que rechaza el contrato es

$$\theta_\phi.$$

Por lo tanto, la disposición a pagar de un vendedor con señal μ por un contrato que especifica una política de revelación $\alpha(\theta)$ es el precio que deja al vendedor indiferente entre aceptar y rechazar, es decir,

$$E_\theta[(\alpha(\theta)\theta + (1 - \alpha(\theta))\theta_\phi) | \mu] - \theta_\phi.$$

Una consecuencia del supuesto de dominación estocástica de primer orden es que la disposición a pagar es una función creciente de la señal μ : vendedores con señales más altas valoran más el servicio de certificación.

El vendedor que sólo tiene información imperfecta sobre la calidad de su producto, puede cometer dos tipos de errores en su decisión de contratar el servicio de certificación. Primero, puede pasar que decida contratar el servicio sobre la base de la señal μ , pero una vez conocido θ desearía no haberlo hecho. Para este vendedor, excesivamente optimista, puede resultar beneficioso que el resultado del test no sea revelado. Pero además, puede ser que la señal indique que no es conveniente contratar al intermediario, cuando la decisión óptima habría sido la contraria, si hubiera conocido perfectamente el valor de θ . Cuanto mayor sea la proporción de este tipo de vendedores excesivamente pesimistas, mayor será el valor que los compradores le asignen a un producto no certificado.

Vamos a ver que existe un equilibrio en el que el beneficio del intermediario, es igual $1/2 - c$. Para eso, supongamos que el intermediario ofrece el siguiente contrato:

$$\{p_0, p(\theta), \alpha(\theta)\} = \{0, \theta, 1\}.$$

Dado que en este contrato la regla de revelación de información es independiente de θ , para que el contrato sea robusto a la renegociación debe pasar que $\theta > \theta_\phi$, $\forall \theta$. Por lo tanto, debe pasar que $\theta_\phi = 0$. Esto significa que los compradores asignan valor 0 al producto cuando el intermediario no revela información.

Dadas esas creencias, con este contrato, cada tipo paga exactamente su máxima disposición a pagar y, por lo tanto, todos los tipos contratan al intermediario. El beneficio esperado del intermediario es

$$E(\theta) - c = \frac{1}{2} - c.$$

En realidad, es posible mostrar que ese es el máximo beneficio que el intermediario puede obtener. La demostración de esta parte es complicada y remitimos al lector interesado a Faure-Grimaud, Peyrache y Quesada (2009). La intuición de la demostración es la siguiente. En principio, el intermediario podría querer subir el precio para dejar algunos tipos con señales muy bajas fuera del mercado. Esto es así, porque la disposición a pagar de un vendedor con señal μ muy cercana al límite inferior es menor al costo del test, c . Sin embargo, dejar esos tipos fuera del mercado aumenta el valor de un producto no certificado (θ_ϕ) y, por lo tanto, reduce la disposición a pagar de todos los demás tipos. Esto implica que es siempre óptimo contratar con algunos tipos cuya disposición a pagar es menor a c . Para el caso de la distribución uniforme, se verifica que es óptimo contratar con todos los tipos. Resumimos este resultado en la siguiente proposición.

Proposición 4 Cualquiera sea $\sigma \geq 0$ el único beneficio de equilibrio robusto a la renegociación es $1/2 - c$. El vendedor contrata al intermediario cualquiera sea su tipo y $\alpha(\theta) = 1, \forall \theta$. Los compradores asignan valor 0 a un producto no certificado.

Es interesante comparar este resultado con el que se obtuvo en la Sección 3. Los resultados son idénticos en cuanto a la distribución del excedente. En efecto, en ambos casos el intermediario se lleva todo el excedente del mercado. Sin embargo, son opuestos en cuanto a la política de revelación de información. Mientras que en la sección anterior es óptimo no revelar ninguna información, en esta toda la información se revela en equilibrio. La diferencia proviene del supuesto de robustez a la renegociación.

Competencia entre intermediarios

El modelo desarrollado hasta ahora supone que el mercado de certificación es monopólico. Por eso se obtiene el resultado de que el intermediario captura todo el excedente del mercado. Lizzeri (1999) muestra que cuando se introduce competencia en el mercado, siempre hay un equilibrio en el que al menos uno de los intermediarios revela toda la información. Es decir, la competencia mejora la información del mercado.

Cuando se incorpora el supuesto de robustez a la renegociación, el resultado es el opuesto. La competencia entre intermediarios reduce la cantidad de información que se revela en equilibrio. En efecto, si introducimos un segundo intermediario y cada intermediario $i = 1, 2$ ofrece un contrato $\{p_0^i, p^i(\theta), \alpha^i(\theta)\}$ robusto a la renegociación, en cualquier equilibrio debe pasar que

$$E_{\theta}(p_0^i + p^i(\theta) | \mu) = c, \quad \forall \mu, \forall i.$$

Esto es así porque la robustez a la renegociación y el hecho de que los compradores no observan la decisión del vendedor de aceptar o no el servicio de certificación hace que los servicios provistos por ambos intermediarios sean idénticos (homogéneos). Se aplica, por lo tanto, un razonamiento del tipo Bertrand que lleva los beneficios de ambos intermediarios a 0, cualquiera sea el tipo del vendedor. De aquí se desprende inmediatamente que existirá siempre un subconjunto de tipos que no se harán certificar porque su disposición a pagar es menor a c . Es decir, existe un valor de la señal, $\tilde{\mu} > -\sigma$ tal que cualquier vendedor con una señal $\mu \in [-\sigma, \tilde{\mu})$ decide vender su producto sin certificar.

El resultado es, en realidad, más fuerte que este. Si $\sigma > c$, en el equilibrio se da que, para algunos de los productos que se hacen certificar, el resultado del test no se revela. Esto

sucede porque cuando la precisión de la señal es baja (σ es grande), la mínima calidad que puede obtener el tipo umbral, $\tilde{\mu}$ es menor que el valor esperado de un producto no certificado i.e., $\tilde{\mu} - \sigma < \theta_{\phi}$.

5. Posibilidad de colusión entre el intermediario y el vendedor

Hasta ahora hemos analizado un modelo en el que el intermediario que decide revelar información, lo hace de manera fidedigna. El supuesto que garantiza esto es que el resultado del test es información sólida. En muchos casos, sin embargo, la colusión entre el intermediario y el vendedor para falsificar información puede ser un problema. El ejemplo más evidente es el de los grandes auditores y las compañías auditadas.⁸

En esta sección vamos a modificar el modelo para poder investigar el problema de la colusión entre el intermediario y el vendedor. Para eso vamos a suponer, como antes, que el test genera evidencia que avala la información. Sin embargo, el intermediario puede ponerse de acuerdo con el vendedor para fraguar evidencia que pueda avalar información falsa. En ese sentido, vamos a decir que la información es *semimaleable* (*semi-soft information*). En segundo lugar, vamos a suponer que una vez realizado el test el intermediario está obligado a revelar alguna información. Hacemos este supuesto para simplificar la estrategia del intermediario, porque sólo debe decidir qué información revelar (la verdadera u otra). Finalmente, para que en algunas ocasiones el intermediario tenga incentivos a revelar la verdadera información, vamos a suponer que este es un juego repetido infinitas veces y el costo de la colusión viene dado por la pérdida de beneficios futuros.

⁸ El caso Enron-Arthur Andersen es sólo un ejemplo.

Presentamos entonces un modelo dinámico de intermediación con infinitos períodos. En cada período t , la calidad del vendedor es θ_t distribuida uniformemente en $[0,1]$, independiente entre períodos. Nos concentraremos en *equilibrios estacionarios*, en los que las estrategias de los jugadores son independientes de t . Por eso, eliminamos el subíndice t a partir de ahora. En cada período, el vendedor conoce perfectamente el valor de θ y el costo del test es 0. Además, los compradores observan la decisión del vendedor de hacerse certificar, pero no observan el contrato entre el vendedor y el intermediario.

Una vez que el intermediario realizó el test, tiene que hacer un anuncio r sobre la calidad del producto. En connivencia con el vendedor y a cambio de un soborno b puede hacer un anuncio falso $r \neq \theta$. Por simplicidad, suponemos que el intermediario tiene todo el poder de negociación en el diseño del contrato de colusión. El contrato de colusión se diseña una vez que el intermediario realizó el test y conoce, por lo tanto, el valor de θ .

En este sentido, el desarrollo temporal del juego se modifica de la siguiente manera. Entre las etapas 4 y 5 se agrega el juego de colusión:

- 4.a) Si el vendedor aceptó el contrato de certificación, el intermediario ofrece un contrato de colusión al vendedor $SC(\theta) = (r(\theta), b(\theta))$ que especifica un anuncio r y un soborno b , que están supeditados al resultado del test.
- 4.b) El vendedor acepta o rechaza el contrato de colusión.
- 4.c) Si el vendedor aceptó el contrato de colusión, el intermediario anuncia r , si no, el intermediario anuncia θ .

Dado que hemos supuesto que para fraguar evidencia se necesita la participación del vendedor, si no se llega a un acuerdo de colusión el intermediario está obligado a revelar la verdad.⁹

Finalmente, suponemos que una vez que los vendedores compran el producto, el valor de θ se revela con certeza. Por lo tanto, los compradores son capaces de identificar ex post si la información revelada por el intermediario fue verdadera o falsa.¹⁰

Igual que en las secciones anteriores, la competencia entre compradores, implica que en equilibrio pagarán un precio por el producto igual al valor esperado de la calidad, dada la información recibida. Es decir, el precio que se pague por el producto será $\beta(l) = E(\theta | l)$, con $l \in [0, 1] \cup \{N\}$. En particular, $l = N$ significa que el vendedor no se hizo certificar y, por lo tanto, el intermediario no hizo ningún anuncio.

Es fácil mostrar que este modelo siempre tiene un equilibrio en el que el intermediario tiene beneficios nulos todos los períodos. Para eso, imaginemos un equilibrio en el que el vendedor, cualquiera sea su tipo, decide no hacerse certificar. Esto significa que $\beta(N) = 1/2$. Entonces, hacerse certificar es una acción que no ocurre en el equilibrio. Supongamos que $\beta(r) = 0, \forall r \in [0, 1]$. Entonces, no existe ningún $p \geq 0$ ni ningún contrato de colusión posible que induzca al vendedor a hacerse certificar. El beneficio del intermediario

⁹ Con esto se elimina la posibilidad de chantaje, en la que el intermediario obtiene un soborno simplemente por amenazar al vendedor con revelar una información muy mala.

¹⁰ El supuesto realmente importante es que exista alguna probabilidad positiva de detectar la colusión. El hecho de que esa probabilidad sea 1, reduce el número de parámetros y simplifica los cálculos.

es nulo en este equilibrio. Vamos a usar este equilibrio como estrategia de castigo para un intermediario que miente.

Supuesto 1 Castigo. Si en el período t los compradores descubren que la calidad es θ y el intermediario hizo un anuncio r tal que $\beta(r) > \theta$, entonces, a partir del período $t + 1$ los compradores se coordinan en el equilibrio de beneficios nulos para siempre.

Suponemos, entonces, que los compradores castigan al intermediario cuando el anuncio realizado les hace pagar por el producto más de su verdadero valor, i.e., $\beta(r(\theta)) > \theta$. Si en cambio $\beta(r(\theta)) \leq \theta$, los compradores siguen en la fase cooperativa, en la que el intermediario puede tener beneficios positivos.

Como lo muestra la existencia del equilibrio con beneficios nulos, las creencias de los compradores son fundamentales para construir los equilibrios. Por eso, vamos a hacer un supuesto que impone alguna estructura a esas creencias para acotar el tipo de equilibrios que pueden ocurrir.

Supuesto 2 Creencias de los compradores. Supongamos que no se ha observado colusión en el pasado. Entonces, (a) dado cualquier anuncio $r \in [0,1]$, la disposición a pagar de los compradores nunca es superior al anuncio, i.e., para cualquier $r \in [0,1]$, $\beta(r) \leq r$. Esto es cierto para anuncios de equilibrio y fuera del equilibrio; (b) los compradores creen cualquier anuncio que no debería hacerse en equilibrio, i.e., si r es un anuncio fuera del equilibrio, entonces $\beta(r) = r$.

El Supuesto 2 es razonable en este modelo. La primera parte simplemente garantiza que el intermediario nunca es castigado cuando revela la verdad. La segunda parte proviene de reconocer que si r es un anuncio fuera del equilibrio, cualquiera sea θ , existe algún otro anuncio que es mejor. Por lo tanto, si el intermediario revela r es porque no llegó a un acuerdo de colusión, en cuyo caso, sólo puede ser la verdad.

Vamos a llamar Π al beneficio esperado de largo plazo del intermediario cuando el juego está en la fase cooperativa. Dado el Supuesto 1 el beneficio de largo plazo en la fase de castigo es 0. Llamamos $\delta \in (0,1)$ a la tasa de preferencia temporal del intermediario, que suponemos conocida por todos los jugadores. Podemos interpretar también a δ como el valor de la reputación, dado que un intermediario con mayor δ se preocupa más por su futuro y está, por lo tanto, más interesado en conservar su reputación.

El equilibrio sin colusión

Vamos a analizar primero un caso en el que el intermediario siempre anuncia la verdad en equilibrio. Es decir, sea $\tilde{\theta}$ el umbral de calidad a partir del cual el vendedor se hace certificar. Entonces, estamos buscando un equilibrio en el que $\beta(r(\theta)) = \theta$, dado cualquier $\theta \in [\tilde{\theta}, 1]$.

Por definición, $\tilde{\theta}$ es el tipo que está indiferente entre hacerse certificar y no hacerse certificar. Si el vendedor se hace certificar, su utilidad es

$$\theta - p,$$

porque el intermediario revela la verdad. Si no se hace certificar, su utilidad es $\beta(N)$. Por lo tanto, $\tilde{\theta}$ es tal que

$$\tilde{\theta} - p = \beta(N).$$

Como los compradores no observan el precio p , las creencias $\beta(N)$ no dependen de la política de tarificación del intermediario.

La demanda total del intermediario es el conjunto de tipos que están por encima de $\tilde{\theta}$. Dada la distribución uniforme, la masa de tipos que contratan al intermediario es $1 - \tilde{\theta}$.

El beneficio esperado de largo plazo del intermediario es

$$\Pi = p(1 - \tilde{\theta}) + \delta\Pi = p(1 - p - \beta(N)) + \delta\Pi.$$

Maximizando el beneficio con respecto al precio obtenemos de la condición de primer orden que

$$(1) \quad p = \frac{1 - \beta(N)}{2}.$$

Además, sabemos que en equilibrio debe darse que

$$\beta(N) = E(\theta | \theta < \tilde{\theta}) = \frac{\tilde{\theta}}{2} = \frac{p + \beta(N)}{2}.$$

Combinando con (1) obtenemos que $p = \beta(N) = \tilde{\theta}/2 = 1/3$. Reemplazando en el beneficio del intermediario nos queda que

$$\Pi = \frac{1}{9(1 - \delta)}.$$

Finalmente, queda por demostrar que el intermediario no tiene incentivos a firmar acuerdos de colusión con el vendedor. Supongamos que la calidad es θ . El mejor contrato de colusión que puede ofrecer el intermediario es $r(\theta) = 1$ acompañado de un soborno $b(\theta) = 1 - \theta$. En efecto, si la colusión implica la pérdida de los beneficios futuros, la mejor colusión posible es la que maximiza el soborno. Por otra parte, como el intermediario tiene todo el poder de negociación, el soborno que va a exigir por un anuncio r es lo

máximo que el vendedor está dispuesto a pagar por ese anuncio. Si acepta la colusión, la utilidad del vendedor es $r - b$. Si no la acepta, el intermediario está obligado a anunciar la verdad y la utilidad del vendedor es θ . Por lo tanto, el máximo soborno que está dispuesto a pagar por un anuncio r es $b(r|\theta) = r - \theta$. Entonces, maximizar el soborno equivale a maximizar r dado cualquier θ .

Si el intermediario ofrece el contrato de colusión, gana el soborno b pero pierde el beneficio de largo plazo desde el período siguiente para siempre. Por lo tanto, el intermediario no quiere ofrecer el contrato de colusión si

$$b(\theta) = 1 - \theta \leq \delta\Pi,$$

o

$$\theta \geq 1 - \delta\Pi = 1 - \frac{\delta}{9(1 - \delta)}.$$

De aquí se desprende que los contratos de colusión más beneficiosos se dan con los tipos más bajos. Vemos entonces que el intermediario no ofrece ningún contrato de colusión en equilibrio si

$$\tilde{\theta} = \frac{2}{3} \geq 1 - \frac{\delta}{9(1 - \delta)} \Leftrightarrow \delta \geq \frac{3}{4}.$$

Proposición 5 Si $\delta \geq 3/4$, existe un equilibrio en el que $r(\theta) = \theta$ para todo $\theta \in [\tilde{\theta}, 1]$. En este equilibrio, $p = 1/3$, $\tilde{\theta} = 2/3$, $\beta(N) = 1/3$ y el beneficio esperado de largo plazo del intermediario es

$$\Pi = \frac{1}{9(1-\delta)}.$$

Cuando δ es suficientemente alto, el costo de la colusión en términos de beneficios perdidos es tan alto que ningún tipo de colusión es aceptable para el intermediario. Ni siquiera el soborno que se puede obtener del vendedor con la calidad más baja ($\tilde{\theta}$) es suficiente para compensar ese costo. Cuando δ disminuye, la tentación de la colusión aumenta, porque el valor de los beneficios perdidos es menor. En particular, para δ suficientemente bajo, no existe ningún equilibrio sin colusión.

Equilibrio con un intermediario impaciente

En esta sección vamos a suponer que $\delta < 3/4$ y por lo tanto un equilibrio como el de la Sección 5.1 no existe. Esto significa que en el equilibrio la probabilidad de colusión será positiva. De acuerdo con la intuición que obtuvimos del caso sin colusión, estamos interesados en encontrar un equilibrio en el que existe otro umbral $\bar{\theta}$ tal que la colusión ocurre sólo con tipos por debajo de ese umbral ($\theta < \bar{\theta}$).¹¹

Lo primero que podemos observar es que el mejor contrato de colusión en este caso es necesariamente distinto del de la sección anterior. En efecto, supongamos que en el equilibrio el intermediario colude con todos los tipos en $[\tilde{\theta}, \bar{\theta}]$ ofreciendo un contrato de colusión $(r, b) = (1, 1 - \theta)$. Entonces, cuando los compradores observan el anuncio $r = 1$, deben creer que la verdadera calidad pertenece al conjunto $[\tilde{\theta}, \bar{\theta}] \cup \{1\}$ y, por lo tanto,

¹¹ Peyrache y Quesada (2011) muestran que estos son los únicos equilibrios posibles.

$$\beta(1) = E(\theta | \theta \in [\tilde{\theta}, \bar{\theta}] \cup \{1\}) = \frac{\bar{\theta} + \tilde{\theta}}{2} < \bar{\theta} \leq 1.$$

Pero esto significa que el soborno que puede pedir es necesariamente menor que $1 - \theta$. Es más, para que el tipo $\bar{\theta}$ quiera aceptar la colusión, tiene que pasar que $\beta(r(\bar{\theta})) > \bar{\theta}$. Es decir, tiene que pasar que, con probabilidad positiva, para algún subconjunto de tipos mayores a $\bar{\theta}$ el anuncio también sea $r(\bar{\theta})$. En particular, podemos demostrar que existe un conjunto $A = [\theta^*, 1] \subset [0, 1]$ tal que el mejor contrato de colusión que puede ofrecer el intermediario es un anuncio $r \in A$ (elegido al azar con una distribución uniforme) y $b = \theta^* - \theta$ y θ^* está definido por la condición

$$\theta^* = E(\theta | \theta \in [\tilde{\theta}, \bar{\theta}] \cup A).$$

La idea es que la estrategia del intermediario consiste en hacer un anuncio $r \in A$ siempre que $\theta \in [\tilde{\theta}, \bar{\theta}] \cup A$. Veamos ahora cómo se completa el equilibrio.

La calidad umbral a partir de la cual el vendedor se hace certificar es nuevamente $\tilde{\theta} = \rho + \beta(N)$. Por otra parte, la máxima calidad con la que el intermediario quiere hacer colusión es $\bar{\theta}$ tal que $b(\bar{\theta}) = \delta\Pi$: el soborno compensa la pérdida del beneficio futuro. Además, sabemos que $b(\theta) = \theta^* - \theta$. Por lo tanto,

$$\bar{\theta} = \theta^* - \delta\Pi.$$

Con esto podemos escribir el beneficio del intermediario como

$$\Pi = p(1 - \tilde{\theta}) + \int_{\tilde{\theta}}^{\bar{\theta}} (\theta^* - \theta) d\theta + (1 - (\bar{\theta} - \tilde{\theta})) \delta \Pi,$$

o

$$(2) \quad \Pi = p(1 - p - \beta(N)) + \int_{p+\beta(N)}^{\theta^* - \delta \Pi} (\theta^* - \theta) d\theta + (1 - (\theta^* - \delta \Pi - p - \beta(N))) \delta \Pi,$$

donde el segundo término representa los ingresos por sobornos y el tercer término es el beneficio futuro, que se obtiene sólo si no hay colusión, es decir, con probabilidad $1 - (\bar{\theta} - \tilde{\theta})$.

Diferenciando con respecto a p e igualando a 0 la derivada obtenemos de la condición de primer orden que

$$(3) \quad p = 1 - \theta^* + \delta \Pi = 1 - \bar{\theta}.$$

Nuevamente, tiene que darse en equilibrio que $\beta(N) = p = \tilde{\theta}/2$.

Finalmente, tenemos que en equilibrio tiene que darse que

$$\theta^* = E(\theta | \theta \in [\tilde{\theta}, \bar{\theta}] \cup [\theta^*, 1]) = \frac{1}{1 - \theta^* + \bar{\theta} - \tilde{\theta}} \left(\int_{\tilde{\theta}}^{\bar{\theta}} \theta d\theta + \int_{\theta^*}^1 \theta d\theta \right),$$

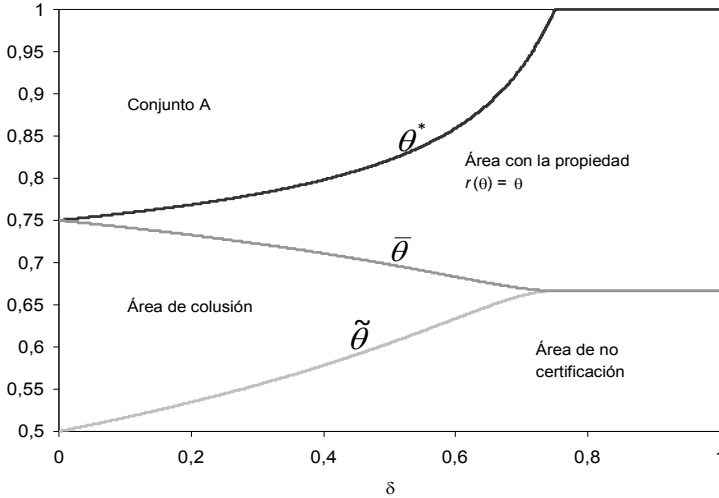
de lo que se obtiene que

$$(4) \quad \theta^* = \frac{1 + (\delta \Pi)^2 - \tilde{\theta}^2}{2(1 - \tilde{\theta})}.$$

El equilibrio se obtiene de la solución de (2), (3) y (4).

Variables de equilibrio como función de la tasa de descuento

Figura 1: Variables de equilibrio como función de δ



Proposición 6 Para todo $\delta < 3/4$, existe un equilibrio en el que el intermediario ofrece un precio $p(\delta)$ y el vendedor se hace certificar si y sólo si $\theta \geq \tilde{\theta}(\delta) = 2p(\delta)$. Existen funciones $\theta^*(\delta)$ y $\bar{\theta}(\delta)$ tales que:

1. $\theta^*(\delta) > \bar{\theta}(\delta) > \tilde{\theta}(\delta)$,
2. $\bar{\theta}(\delta) = \theta^*(\delta) - \delta\Pi(\delta)$,
3. $p(\delta) = 1 - \bar{\theta}(\delta)$,
4. $\theta^*(\delta)$ es creciente en δ y $\bar{\theta}(\delta)$ es decreciente en δ , con $\theta^*(3/4) = 1$, $\bar{\theta}(3/4) = \tilde{\theta}(3/4) = 2/3$ y $\theta^*(0) = \bar{\theta}(0) = 3/4$.

5. *La representación numérica de la solución se presenta en la Figura 1.*

La Figura 1 resume las variables de equilibrio como función del parámetro de preferencia temporal del intermediario, δ . En primer lugar, se observa que cuando δ es bajo los compradores no confían plenamente en el intermediario cuando hace anuncios altos. Esto es porque saben que, en algunas ocasiones, ese anuncio va a ser el resultado de un acuerdo de colusión con el vendedor. Por eso, reducen la disposición a pagar por el producto cuando observan un anuncio alto: $\beta(r) = \theta^* < r \forall r \in (\theta^*, 1]$. Por el contrario, los compradores confían en el intermediario cuando el anuncio es moderado: $\beta(r) = r \forall r \in [0, \theta^*]$. Una forma de interpretar este resultado es que los compradores reaccionan más a las malas noticias que a las buenas.

Segundo, el precio que cobra el intermediario por el servicio de certificación y la calidad umbral a partir de la cual el vendedor elige la certificación son funciones crecientes de δ . Es decir, un intermediario con δ alto pone un precio alto al servicio de certificación como una forma de atarse las manos, para no caer en la tentación de contratar con vendedores de calidad muy baja, con los que las oportunidades de colusión serían beneficiosas.

Por último, observamos que la probabilidad de que ocurra colusión en el equilibrio (la diferencia entre $\bar{\theta}$ y $\tilde{\theta}$) decrece a medida que aumenta δ . Esto concuerda con la idea de que un intermediario más paciente se ve menos tentado de entrar en acuerdos de colusión. Sin embargo, observamos también que la probabilidad de colusión es siempre menor que 1, incluso cuando $\delta = 0$. Esto es así por dos razones. La primera es que aun cuando $\delta = 0$ el intermediario no quiere contratar con un vendedor de calidad cercana a 0

($\tilde{\theta}(0)=1/2$). Por lo tanto, existe una probabilidad positiva de que el vendedor no se haga certificar y el contrato de colusión en ese caso no existe. La segunda razón es la más interesante. Como $\theta^* < 1$ (y decrece cuando δ disminuye), existe una probabilidad positiva de que la calidad del vendedor que contrata al intermediario esté por encima de θ^* . Ningún contrato de colusión es beneficioso en este caso, porque no existe ningún anuncio que induzca a los compradores a pagar por el producto un precio superior a θ^* . Esto significa que es posible que un intermediario impaciente sobreviva en el mercado por varios períodos. En particular, podemos calcular la esperanza de vida del intermediario como función de δ .

Llamamos T al último período en el que el juego se encuentra en la fase cooperativa. La probabilidad de que el intermediario desaparezca en algún período t es igual a la probabilidad de que sobreviva hasta el período t (no colude en ningún momento hasta $t - 1$) y que coluda en el período t :

$$\Pr(T = t | \delta) = \left(1 - \left(\bar{\theta}(\delta) - \tilde{\theta}(\delta)\right)\right)^{t-1} \left(\bar{\theta}(\delta) - \tilde{\theta}(\delta)\right)$$

Por lo tanto, la esperanza de vida del intermediario es

$$T^*(\delta) = E(T | \delta) = \sum_{t=1}^{\infty} t \left(1 - \left(\bar{\theta}(\delta) - \tilde{\theta}(\delta)\right)\right)^{t-1} \left(\bar{\theta}(\delta) - \tilde{\theta}(\delta)\right) = \frac{1}{\bar{\theta}(\delta) - \tilde{\theta}(\delta)} \geq 4,$$

creciente en δ .

Esto nos muestra que, a pesar de que los compradores conocen el grado de impaciencia del intermediario, éste sobrevive durante un tiempo en el mercado antes de entrar en un acuerdo de colusión. La esperanza de vida se alarga, por

supuesto, se reduce cuando el intermediario se vuelve más paciente.

6. Extensiones y cuestiones pendientes

El objetivo de este trabajo fue introducir el mercado de los intermediarios de certificación, como instituciones de mercado tendientes a reducir las asimetrías de información entre compradores y vendedores. En este sentido, presentamos tres modelos que analizan distintos aspectos de estas instituciones. En esta sección vamos a presentar otros aspectos que por cuestiones de espacio hemos dejado de lado.

Tecnología de testeo imperfecta. Los tres modelos presentados en este trabajo comparten el supuesto de que la tecnología del intermediario le permite conocer con certeza la calidad del producto. Gu (2008) analiza los tipos de equilibrios que podrían existir si la tecnología de testeo fuera imperfecta. Sin embargo, ese modelo no es estrictamente comparable con los presentados aquí, ya que reduce la estrategia de revelación de información del intermediario a decir si la calidad es superior o no a un estándar dado. La introducción de ruido en la tecnología del intermediario, sin embargo, introduce nuevas cuestiones de análisis. Una de ellas consiste en entender cuáles son los incentivos a invertir en mejorar la tecnología de testeo (reducir el ruido). Es posible que el beneficio extra por mejorar la tecnología sea mayor cuando haya competencia entre intermediarios, ya que servirá como un elemento diferenciador de producto, eliminando el tipo de competencia a la Bertrand que reduce los beneficios a cero. Segundo, supongamos que el resultado del test también resulta informativo para el vendedor (porque su propia información es imprecisa). Con tecnología perfecta, no existe ninguna razón por la cual el vendedor pudiera querer pagar por un segundo test de calidad. Esto ya no es cierto con tecnología imperfecta. Si la tecnología del

intermediario es ruidosa, el vendedor podría querer pedir una “segunda opinión”, de un intermediario competidor.

Obligación de revelar información. Los dos primeros modelos desarrollados en este trabajo suponen que el intermediario puede decidir, como parte de su contrato de certificación, no revelar la información obtenida a partir del test. Podría pensarse en una regulación que obligara a los intermediarios a publicar los resultados de los tests. Dado que ocultar la información es en muchos casos óptimo, este tipo de regulación va a tener efectos sobre la política de tarificación del servicio de certificación. En particular, si el intermediario está obligado a revelar el resultado del test, querrá ofrecer contratos que sean aceptados sólo por vendedores de productos de alta calidad. En efecto, los vendedores de productos de baja calidad sólo tienen interés en contratar al intermediario si eso les permite confundirse con otros productos de calidad más alta. Por lo tanto, la cantidad de productos certificados se reduce si se impone una obligación de revelación de información. Faure-Grimaud, Peyrache y Quesada (2009) muestran este resultado para el caso en el que el equilibrio predice que una parte de la información no se revela.

Auditoría (certificación) obligatoria. En muchos casos, las firmas están obligadas a contratar un servicio de certificación. Esto se da especialmente en el caso de las auditorías (por ejemplo, como requisito para cotizar en la bolsa de comercio). En los modelos presentados en este trabajo, suponimos que el vendedor podía decidir no contratar al intermediario si lo consideraba conveniente. ¿Cómo se modificarían los resultados presentados si agregáramos esta restricción? Está claro que nada cambia en los casos en los que el vendedor siempre elige certificarse en el equilibrio. Los efectos más interesantes de este tipo de regulación los encontramos en el contexto de la Sección 5 en el que el intermediario y el

vendedor pueden tener incentivos a coludir. En efecto, vemos que en ese caso los productores de menor calidad no contratan el servicio de certificación en el equilibrio. Más aún, el conjunto que queda fuera del mercado de certificación es mayor cuanto mayor es el valor de la reputación. El intermediario usa la posibilidad de excluir a los tipos más bajos como herramienta para reducir la rentabilidad de los acuerdos de colusión. Por lo tanto, si introducimos una regulación de certificación obligatoria, productos de menor calidad se harán certificar y más acuerdos de colusión resultarán provechosos. Esto implica que la obligatoriedad de la certificación aumenta la probabilidad de colusión en el mercado. Peyrache y Quesada (2011) muestran cómo aumenta la probabilidad de colusión cuando se introduce esta regulación.

7. Referencias bibliográficas

- Akerlof, G. (1970): "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84, 488-500.
- Albano, G.-L., y A. Lizzeri (2001): "Strategic Certification and Provision of Quality," *International Economic Review*, 42, 267-283.
- Bénabou, R., y G. Laroque (1992): "Using Privileged Information to Manipulate Markets: Insiders, Gurus and Credibility," *Quarterly Journal of Economics*, 107, 921-958.
- Biglaiser, G. (1993): "Middlemen as Experts," *Rand Journal of Economics*, 24, 212-223.
- Biglaiser, G., y J. Friedman (1994): "Middlemen as Guarantors of Quality," *International Journal of Industrial Organization*, 12, 509-531.
- Crawford, V., y J. Sobel (1982): "Strategic Information Transmission," *Econometrica*, 50, 1431-1451.

- Faure-Grimaud, A., E. Peyrache, y L. Quesada (2009): "The Ownership of Ratings," *Rand Journal of Economics*, 40, 234-257.
- Gu, Y. (2008): "Imperfect Certification," Ruhr Economic Papers # 78.
- Lizzeri, A. (1999): "Information Revelation and Certification Intermediaries," *Rand Journal of Economics*, 30, 214-231.
- Mathios, A. (2000): "The Impact of Mandatory Disclosure Laws on Product Choices: An Analysis of the Salad Dressing Market", *Journal of Law and Economics*, 43, 651-667.
- Matthews, S. and A. Postlewaite (1985): "Quality Testing and Disclosure," *Rand Journal of Economics*, 16, 328-340.
- Meggison, W., y K. Weiss (1991): "Venture Capitalist Certification in Initial Public Offerings," *Journal of Finance*, 46, 879-903.
- Morris, S. (2001): "Political Correctness," *Journal of Political Economy*, 109, 231-265.
- Okuno-Fujiwara, M., Postlewaite, A. y Suzumura, K. (1990): "Strategic Information Revelation," *Review of Economic Studies*, 57, 25-47.
- Peyrache, E., y L. Quesada (2011): "Intermediaries, Credibility and Incentives to Collude," *Journal of Economics and Management Strategy*, forthcoming.
- Rubinstein, A., y A. Wolinsky (1987): "Middlemen," *Quarterly Journal of Economics*, 102, 581-593.
- Strausz, R. (2005): "Honest Certification and the Threat of Capture," *International Journal of Industrial Organization*, 23, 45-62.

Verrecchia, R. (1983): “Discretionary Disclosure,” *Journal of Accounting and Economics*, 5, 179-194.

Verrecchia, R. (2001): “Essays on disclosure,” *Journal of Accounting and Economics*, 32, 97-180.

