ADOLFO FIGUEROA

LA SOCIEDAD SIGMA:

una teoría del desarrollo económico

Pontificia Universidad Católica del Perú
Fondo Editorial 2003

FONDO DE CULTURA ECONÓMICA

Adolfo Figueroa obtuvo su doctorado en Economía (*Ph. D. in Economics*) en la Universidad de Vanderbilt, Nashville, Estados Unidos. Actualmente, es decano de la Facultad de Ciencias Sociales de la Pontificia Universidad Católica del Perú y profesor principal del Departamento de Economía en la misma casa de estudios. Ha sido profesor visitante en la Universidad de Illinois, la Universidad de Notre Dame, la Universidad de Texas y la Universidad de Wisconsin en los Estados Unidos, así como en la Universidad de Oxford en Inglaterra. Además de su labor académica y docente, se ha desempeñado como consultor de varias organizaciones internacionales como el Banco Mundial, la OIT, la FAO, el FIDA y el BID. Producto de sus investigaciones en el campo del desarrollo económico, ha publicado diversos artículos y varios libros. Entre sus obras más importantes destaca *Capitalist Development and the Peasant Economy in Peru*, publicado por la prestigiosa editorial Cambridge University Press de Inglaterra.

LA SOCIEDAD SIGMA:
una teoría del desarrollo económico

LA SOCIEDAD SIGMA:
UNA TEORÍA DEL DESARROLLO ECONÓMICO

Adolfo Figueroa

Pontificia Universidad Católica del Perú
Fondo Editorial

Fondo de Cultura Económica

La sociedad Sigma: una teoría del desarrollo económico

Primera edición: octubre del 2003

Tiraje: 750 ejemplares

© 2003, Adolfo Figueroa

© 2003 de esta edición por Fondo Editorial de la Pontificia Universidad Católica del Perú.
Plaza Francia 1164, Cercado, Lima - Perú
Telefáx: 330-7405; 330-7410; 330-7411
E-mail: feditor@pucp.edu.pe

© 2003 de esta edición por Fondo de Cultura Económica
Carretera Picacho Ajusco 277; 14200, México D.F.

© 2003 de esta edición por Fondo de Cultura Económica del Perú
Jr. Berlín 238, Lima 18

www.fondodeculturaeconomica.com

Diseño de cubierta: Magaly Sánchez
Correción de estilo: Mario Naranjo
Cuidado de la edición: Óscar Hidalgo

Prohibida la reproducción de este libro por cualquier medio total o parcialmente, sin permiso expreso de los editores.

Derechos reservados

ISBN: 9972-42-597-5
Hecho el Depósito Legal N.º 1501052003-5025

Impreso en Perú - Printed in Peru

A la memoria de mi profesor
Nicholas Georgescu-Roegen

Contenido

INTRODUCCIÓN 19

1. LA ECONOMÍA COMO CIENCIA SOCIAL
   - 1.1. La ciencia social como el estudio de procesos complejos 30
   - 1.2. La metodología alfa-beta 34
   - 1.3. La metodología alfa-beta en la economía 51
   - 1.4. El estudio científico del capitalismo contemporáneo 64

2. LAS ECONOMÍAS CONVENCIONALES
   - 2.1. La economía neoclásica 67
   - 2.2. La economía clásica 88
   - 2.3. La economía keynesiana 101

3. LA SOCIEDAD ÉPSILON
   - 3.1. Una sociedad de clases sociales 115
   - 3.2. La naturaleza del mercado de trabajo 116
   - 3.3. Un modelo estático de la economía épsilon 126
   - 3.4. Equilibrio general 131
   - 3.5. Proposiciones beta 133
   - 3.6. Consistencia empírica: los países del primer mundo 141

4. UN MODELO ABIERTO DE ÉPSILON
   - 4.1. Un modelo con tipo de cambio flexible 143
   - 4.2. El equilibrio general 149
   - 4.3. Proposiciones beta 153
   - 4.4. Un modelo con mercado de crédito bancario 158
   - 4.5. Resumen de la economía épsilon estática y abierta 166

5. LA SOCIEDAD OMEGA
5.1. Una sociedad sobrepoblada 171
5.2. Un modelo de la economía omega 174
5.3. Equilibrio general 181
5.4. Proposiciones beta 188
5.5. Consistencia empírica: los países del tercer mundo 193

6. LA SOCIEDAD SIGMA
6.1. Una sociedad heterogénea y jerárquica 195
6.2. Un modelo estático de la economía sigma 198
6.3. El equilibrio general 200
6.4. Proposiciones beta 205
6.5. La naturaleza del equilibrio en la sociedad sigma 208
6.6. Consistencia empírica: los países latinoamericanos 217
6.7. Conclusiones 228

7. MODELOS DINÁMICOS
7.1. Un modelo dinámico para la economía épsilon 232
7.2. Progreso tecnológico 241
7.3. Un modelo dinámico de la economía omega 247
7.4. Un modelo dinámico de la economía sigma 249
7.5. Equilibrio y reproducción social 251

8. BIENES PÚBLICOS E INVERSIÓN PRIVADA
8.1. Tolerancia social a la desigualdad 260
8.2. Equilibrio general con desorden social 266
8.3. Proposiciones beta 269
8.4. Bienes públicos en la función de producción 270
8.5. Bienes públicos en la inversión privada 272
8.6. Consistencia empírica: la economía mundial 278
8.7. Desigualdad, inversión y competitividad internacional 283

9. UNA TEORÍA GENERAL DEL CAPITALISMO
9.1. Teoría de la inclusión-exclusión 291
9.2. La desigualdad de las naciones 299

9.3. Orden social y desigualdad 307
9.4. El problema dual de la convergencia 312
9.5. El desarrollo económico y su evolución 322
9.6. Las implicaciones de política para las sociedades del tipo sigma 325

APÉNDICE:
AVERSIÓN AL RIESGO FRENTE A AVERSIÓN A JUEGOS RIESGOSOS

1. Teoría del ingreso esperado 334
2. Teoría de Bernoulli 335
3. Teoría de la utilidad 336
4. Teoría de la media-varianza 337
5. ¿Se puede utilizar probabilidades como variable exógena? 339
6. El punto de desastre de Hicks 342
7. Una teoría lexicográfica de la elección de portafolio 344

BIBLIOGRAFÍA 347

Lista de figuras

1.1	Representación diagramática de un proceso	31
1.2	La metodología alfa-beta	40
2.1	Equilibrio general neoclásico	81
2.2	Equilibrio general clásico	97
2.3	Equilibrio general keynesiano	108
3.1	Eficiencia y esfuerzo laboral	124
3.2	Equilibrio general con desempleo	134
3.3	Efecto de la acumulación de capital sobre el equilibrio en el mercado laboral	138
4.1	Equilibrio general en la economía épsilon	148
4.2	Equilibrio en los mercados monetario y laboral	153
4.3	Equilibrio en el mercado de crédito	159
4.4	Curva de Lorenz sobre la distribución de ingresos en la economía epsilón	168
5.1	Equilibrio general en la economía omega	179
5.2	Curva de Lorenz sobre la distribución de ingresos en la economía omega	187
6.1	Equilibrio general en la economía sigma	202
6.2	Curva de Lorenz sobre la distribución de ingresos en la economía sigma	204
7.1	Efecto de la oferta monetaria sobre la trayectoria del producto	237
7.2	Efecto de la inversión sobre la curva de demanda de trabajo	243
8.1	Región de la tolerancia social a la desigualdad	262
9.1	Tipos de equilibrio dinámico	313
9.2	Desigualdad y niveles de ingreso por regiones del mundo	321
A.1	Equilibrio en la teoría de la media-varianza	339

A.2 La teoría de Hicks 343
A.3 Aversión a juegos riesgosos 345

Lista de cuadros

2.1 Matriz de proposiciones beta de un modelo neoclásico estático 85
2.2 Matriz de proposiciones beta de un modelo clásico estático 99
2.3 Matriz de proposiciones beta de un modelo keynesiano estático 112
3.1 Matriz de proposiciones beta de un modelo épsilon estático y cerrado 140
4.1 Matriz de proposiciones beta de un modelo épsilon estático y abierto 156
5.1 Matriz de proposiciones beta de un modelo omega estático ($D_h < L^*$) 190
6.1 Matriz de proposiciones beta de un modelo sigma estático ($D_h < L^*$) 206
6.2 América Latina: desigualdad en ocho países en años seleccionados en el periodo 1950-1993 219
8.1 América Latina: inversión directa extranjera 282
8.2 América Latina: participación en el comercio mundial (porcentajes) 285
8.3 América Latina: estructura de exportaciones (porcentajes) 286
9.1 Características de las economías épsilon, omega y sigma 296
9.2 Desigualdad por regiones del mundo, 1950-1995 302
9.3 Desigualdad por regiones del mundo, por décadas, 1959-1995 304
9.4 Coeficientes de Gini por países seleccionados, 1960-1995 305
A.1 Tabla de contingencia 334

Símbolos utilizados

α	Proposiciones teóricas
β	Proposiciones con contenido empírico, derivados de α
δ	Distribución inicial de activos entre individuos
λ	Indice de esfuerzo laboral en la firma
B	Bien producido domésticamente
C	Bien importado
D	Grado de desigualdad en la distribución del ingreso nacional
D_i	Cantidad demandada del bien o servicio i
E	Divisa
FI	Inversión directa extranjera
H	Servicios laborales
I	Inversión privada neta y real
K_b	*Stock* de capital físico
K_b^i	*Stock* de capital físico en la forma de infraestructura pública
K_h	*Stock* de capital humano
L	Cantidad de personas
L^*	Cantidad máxima de empleo asalariado
M	Dinero doméstico
n	Coeficiente de importación
P	Beneficios o ganancias totales de las firmas, en términos reales
P'	Beneficios o ganancias totales de las firmas, en términos nominales
P_b	Precio nominal del bien B en el mercado doméstico
P_b^*	Precio internacional del bien B, en unidades de la divisa
P_c	Precio nominal del bien C en el mercado doméstico
P_c^*	Precio internacional del bien C, en unidades de la divisa

P_e Tipo de cambio nominal
P_h Tasa de salario nominal
Q_b Cantidad producida domésticamente del bien B en el sector capitalista
r Tasa de interés doméstica y nominal
r^* Tasa de interés internacional y nominal
S Grado de orden social
S_i Cantidad ofrecida del bien o servicio i
s Tasa de prima (precio) en el mercado de seguros
t Tiempo como variable discreta
U Cantidad total de desempleo
u Tasa de desempleo
u^* Umbral de la tasa de desempleo de eficiencia
V Producto total en el sector de subsistencia
v Producto medio en el sector de subsistencia
v' Producto marginal en el sector de subsistencia
W Masa salarial real
w Tasa salarial real
w^* Umbral de la tasa salarial real de eficiencia
X_b Cantidad exportada del bien B
Y Ingreso nacional real
Z *Stock* de bienes públicos
z^* Coeficiente de los términos de intercambio

Agradecimientos

MUCHAS PERSONAS ME HAN AYUDADO a producir este libro. El primer manuscrito lo completé y enseñé durante mi estadía como profesor visitante en la Universidad de Texas en Austin en 1997. Luego produje otras versiones que utilicé para enseñar en la Facultad de Ciencias Sociales de la Pontificia Universidad Católica del Perú, y también durante mi estadía como profesor visitante en la Universidad de Wisconsin en Madison en 2001. Debo agradecer a mis estudiantes en todas estas universidades, quienes con sus preguntas contribuyeron al desarrollo de la teoría que ahora presento.

En esta alternancia entre universidades, mis manuscritos eran escritos parte en inglés y parte en castellano. Gracias a la generosidad de un grupo de estudiantes de Economía de mi universidad —que hizo las traducciones al castellano—, he podido producir esta versión castellana del libro. Quiero, en consecuencia, agradecer a Claudia Barriga, Gabriela Cuadra, David Florián, Sofía León, Luis Orezzoli, Jorge Trelles y Ana María Whittembury.

Graciela Zevallos colaboró en este proyecto como asistente de investigación. Beatriz Gonzales y Frida Beltrán me ayudaron a producir innumerables versiones de este libro. Les quiero expresar mi más profunda gratitud.

Varios colegas de la Pontificia Universidad Católica del Perú leyeron partes del manuscrito en sus distintas etapas de desarrollo y quiero expresarles mi agradecimiento. Ellos son Paul Phumpiu, Ramón García-Cobián y Jeanine Anderson. Ciertamente a mis colegas del Departamento de Economía les debo el ambiente de continua confrontación intelectual tan necesario para proyectos como el que ahora publico. Así mismo, quisiera

agradecer a mis colegas Michael Carter (Universidad de Wisconsin), Jaime Ros (Universidad de Notre Dame) y Rosemary Thorp (Universidad de Oxford) por sus aportes en medio de discusiones muy productivas y siempre amigables.

Introducción

ALGUNOS HECHOS BÁSICOS de la economía mundial contemporánea permanecen sin explicación. Uno de ellos se refiere a las disparidades persistentes en el nivel de desarrollo económico entre los países del primer mundo y los del tercer mundo. Los estudios empíricos coinciden en mostrar que las brechas en el ingreso per cápita, un componente del desarrollo económico, entre los dos grupos de países no se están cerrando (cf. Barro y Sala-i-Martin 1995). Este problema se conoce en la literatura actual como «el problema de la convergencia». En cuanto a la igualdad en la distribución del ingreso, el otro componente del desarrollo, la evidencia empírica es también clara en mostrar que el grado de desigualdad en los países del tercer mundo es mayor que en los países del primer mundo, y que esta diferencia es persistente (cf. Li, Squire y Zou 1998). Este caso también constituye un problema de convergencia, pero no ha recibido mayor atención en la literatura. Existen entonces dos problemas de convergencia en la economía mundial que necesitan ser explicados.

La falta de convergencia en los niveles de ingresos se ha intentado explicar dentro de los modelos de crecimiento económico de la teoría neoclásica. Pero, como se mostrará en este estudio, esta teoría tiene problemas para ofrecer tal explicación. La rama de la teoría neoclásica que se ocupa de estudiar el tercer mundo se denomina «economía del desarrollo». Sin embargo, los campos de la economía del desarrollo y de la teoría del crecimiento son independientes. Recientemente se han producido trabajos que han logrado integrar los modelos de crecimiento con los modelos de la teoría del desarrollo (cf. Ros

2000). Dichos trabajos constituyen contribuciones importantes a la unidad del conocimiento en la ciencia económica. Sin embargo, el problema dual de la convergencia señalado arriba —la persistente disparidad tanto en el nivel del ingreso como en la desigualdad— está todavía pendiente de explicación.

Las acciones de política conducentes al desarrollo económico están, de manera explícita o implícita, basadas en el actual paradigma de la ciencia económica. Un criterio sencillo para reconocer el paradigma consiste en identificar la teoría que se enseña en los textos universitarios. Y el actual paradigma en la ciencia económica lo constituye la teoría neoclásica. Basadas en esta teoría, las políticas de desarrollo aplicadas al tercer mundo no han funcionado. La persistencia de las disparidades señaladas es la prueba de ello.

El problema dual de la convergencia ocurre en un periodo donde la revolución de la información ha generado un mundo más globalizado en la economía mundial. De esta manera, el pronunciado grado de desigualdad en la economía mundial es más flagrante que nunca, y las consecuentes tensiones sociales son más agudas. Por lo tanto, el problema dual de la convergencia puede ser considerado el problema social fundamental de nuestro tiempo.

Este estudio tiene el objetivo de desarrollar una nueva teoría que pueda dar cuenta del problema dual de la convergencia. Es una teoría general del desarrollo del capitalismo. Esta teoría parte del supuesto de que existen varios tipos de sociedades capitalistas. Estas sociedades se diferencian en las condiciones iniciales bajo las cuales nacieron al capitalismo. Entre las condiciones iniciales se incluyen no solo la dotación inicial de los factores productivos o activos económicos (capital, tierra y trabajo) y su distribución entre los individuos, sino, también, la existencia de una dotación inicial de activos sociales y su distribución entre los individuos. El supuesto es que existe una desigualdad inicial en la dotación de los activos económicos y sociales en todos los tipos de sociedades capitalistas, pero que los grados de des-

igualdad varían y dan lugar a diferentes tipos de sociedades capitalistas.

Otro supuesto de la teoría es que el capitalismo es un sistema que opera con mecanismos de inclusión y exclusión de los individuos en el proceso económico y político. Pero estos mecanismos operan de manera distinta de acuerdo con el tipo de sociedad, según sean sus condiciones iniciales. Dos preguntas que se derivan de estos supuesto son las siguientes: ¿cómo funcionan estos distintos tipos de sociedades?, ¿tienden ellos de manera natural a la convergencia? En este estudio, se busca darles una respuesta.

La respuesta general propuesta en este estudio es una teoría. Entonces debe cumplir con los requisitos básicos de la epistemología. Se prestará especial atención a dos de esos requerimientos. En primer lugar, esta teoría general debe conducirnos a la unidad del conocimiento. No puede darnos explicaciones separadas e inconsistentes entre ellas, según se trate de la microeconomía o la macroeconomía, del corto plazo o del largo plazo, de los países del primer mundo o los del tercer mundo. En segundo lugar, esta teoría debe producir proposiciones que sean empíricamente refutables.

En los siguientes párrafos pasamos a resumir el contenido del libro.

Debido a que el presente trabajo presenta una nueva teoría es importante que las posibles fuentes de desacuerdo con sus conclusiones sean identificables. Una de las fuentes de desacuerdo en la ciencia es la metodología que se usa. Luego, este libro presenta de manera explícita la teoría del conocimiento que utilizará. La teoría del conocimiento es una teoría normativa que se necesita para desarrollar teorías positivas (con capacidad para explicar la realidad). En el capítulo 1, se presenta un modelo particular de la teoría del conocimiento desarrollada por el filósofo de la ciencia Karl Popper, bajo la denominación de «metodología alfa-beta». Este modelo también se inspira en los trabajos de dos grandes economistas: Nicholas Georgescu-Roegen y Paul Samuelson.

En la ciencia económica actual coexisten varias teorías. Si nos guiáramos por el contenido de los textos universitarios, la teoría neoclásica sería definida como el paradigma actual en la ciencia económica. Pero la teoría keynesiana y, en menor medida, la teoría clásica (basada en los trabajos de Marx y Ricardo) también se utilizan en las discusiones actuales. Se puede, entonces, llamar a estas tres teorías la *economía estándar* actual. La estructura básica de estas teorías, a la luz de los criterios epistemológicos establecidos, se presenta en el capítulo 2.

La conclusión es que estas tres teorías tienen problemas para explicar el desarrollo del capitalismo. En particular, tienen problemas para explicar el funcionamiento de los mercados laborales. La existencia y persistencia del desempleo en el primer mundo no puede ser explicada por estas teorías. En realidad, según estas teorías, la economía podría operar a pleno empleo, pues el desempleo no juega papel alguno en el funcionamiento del sistema. La economía estándar tampoco puede explicar la coexistencia del desempleo y subempleo que se observa en los países del tercer mundo.

Debido a los problemas que tiene la economía estándar, se propone una nueva teoría general del desarrollo del capitalismo. Esta teoría general se inicia con la construcción de tres sociedades abstractas. Para que no quede duda alguna de que se trata de sociedades abstractas se les ha puesto nombres de letras griegas: épsilon, omega y sigma. En términos de sus dotaciones iniciales de factores, épsilon está altamente dotada de capital con relación al trabajo, sigma está pobremente dotada de capital y omega está en una situación intermedia. En términos de la desigualdad inicial, existe desigualdad en los activos económicos pero igualdad en los activos sociales en épsilon y omega; en cambio, en sigma existe también desigualdad en los activos sociales. Sigma «nació» al capitalismo como una sociedad socialmente heterogénea y jerárquica, mientras que las otras dos «nacieron» como sociedades socialmente homogéneas.

En las tres sociedades, existen mecanismos de inclusión de las personas a la sociedad a través del sistema de mercado y del sistema político. También existen mecanismos de exclusión. Algunas personas son excluidas de los *mercados básicos,* definidos como aquellos que juegan un papel central en la producción y reproducción de la desigualdad (mercados del trabajo, crédito y seguros). La exclusión del mercado toma lugar en las tres sociedades. También algunas personas son excluidas del sistema político. A dicha exclusión se le llama exclusión social y toma lugar solo en la sociedad sigma.

Dada esas diferencias en las condiciones iniciales, se muestra que esas sociedades tienen trayectorias distintas en el proceso de desarrollo económico. Épsilon es más desarrollada, sigma es la menos desarrollada y omega se ubica en el medio. No hay nada automático ni espontáneo que lleve a la sociedad sigma a alcanzar a la sociedad épsilon; en cambio, la sociedad omega sí puede hacerlo. En suma, la teoría general predice falta de convergencia generalizada. En el proceso de desarrollo económico de los países, las condiciones iniciales son importantes; es decir, la historia cuenta. Esta predicción es consistente con los hechos estilizados de la economía mundial presentados arriba.

A esa conclusión se llega haciendo iteraciones entre teoría económica y evidencia empírica sobre cada una de las sociedades. Se inicia la iteración utilizando modelos estáticos. Las tres sociedades funcionan con mecanismos diferenciados de inclusión y de exclusión de trabajadores al proceso económico. Aunque en las tres sociedades existe exclusión económica, en sigma existe, además, exclusión social. La hipótesis empírica es que los países del primer mundo se parecen a la sociedad épsilon; que los países del tercer mundo, que tienen el legado de una dominación colonial, se parecen a la sociedad sigma; y que el resto de países del tercer mundo se parecen a la sociedad omega. Esta hipótesis es corroborada empíricamente. Los capítulos 3 al 6 presentan estos resultados.

Debido a que se quiere explicar las diferencias en la trayectoria de desarrollo de cada sociedad, se utilizan modelos dinámicos en los que se busca explicar la inversión privada. Estos modelos muestran el proceso de desarrollo endógeno de cada una de las sociedades. Se desarrolla una teoría de la inversión privada, en la que esta inversión depende del *stock* de bienes públicos. Entre los bienes públicos se incluye el orden social; y como el orden social depende de la desigualdad inicial, también la trayectoria del crecimiento económico depende de las condiciones iniciales de las sociedades. Por otro lado, el grado de desigualdad en la distribución de los *flujos* de ingreso depende de la desigualdad inicial en la distribución de los *stocks* de activos económicos y sociales. La desigualdad tiende a perpetuarse en cada sociedad, aunque a distintos grados de desigualdad. Los capítulos 7 y 8 presentan estos resultados.

La teoría general es el resultado de una consolidación de las hipótesis teóricas que se utilizaron para explicar la producción y la distribución en cada una de las sociedades capitalistas estudiadas. Esta teoría general intenta explicar el comportamiento del conjunto de las sociedades capitalistas; en consecuencia, aspira a constituirse en una teoría general del desarrollo del capitalismo. En efecto, las predicciones de esta teoría general son consistentes con los datos básicos de la economía mundial. Así, se puede explicar por qué los países mantienen casi invariable su grado de desigualdad absoluta y relativa. La desigualdad es una característica estructural de los países. Utilizando una distinción que el castellano nos permite hacer, es más propio decir que, por ejemplo, «el Perú es desigual» que decir que «el Perú *está* desigual». Se explica también el orden en el nivel de ingresos de los países. Así, la teoría general explica por qué no hay convergencia en el sentido dual, ni en el nivel de ingreso ni en la desigualdad, que es el hecho estilizado más importante de la economía mundial actual. Estos resultados se presentan en el capítulo 9.

El hecho de que persistan las disparidades entre los países del primer mundo y los del tercer mundo constituye una paradoja. Las brechas que existen en la tecnología entre estos dos grupos de países crea muchas posibilidades de crecimiento para los países del tercer mundo, pues es más fácil adoptar que crear nuevas tecnologías. Esta paradoja se explica, según la teoría general, por las condiciones iniciales de los países. Entre estas, la desigualdad inicial juega un papel central. La teoría general, en efecto, le da sustento a la proposición de que, en el proceso de desarrollo económico de los países capitalistas, su historia cuenta.

Según la teoría general, las sociedades capitalistas funcionan con mecanismos de inclusión y exclusión de la población en el proceso económico. El grado de desigualdad de una sociedad se debe no solo a los problemas de inclusión (como la explotación a través del intercambio del mercado) sino, también, a los problemas de exclusión. Aún más, cuanto mayor sea el grado de exclusión, mayor es el grado de desigualdad de la sociedad. Esta predicción es consistente con la disparidad persistente en el grado de desigualdad del ingreso entre los países del primer mundo y los del tercer mundo. Por contraste, en la teoría neoclásica la desigualdad es el resultado de los mecanismos de inclusión solamente. La teoría general puede, en consecuencia, denominarse *teoría de la inclusión-exclusión.*

La conclusión de que el proceso de desarrollo depende de las condiciones iniciales no significa determinismo histórico. La teoría general contiene variables exógenas, y dichas variables pueden ser utilizadas como instrumentos de política. Ciertamente, implica políticas de desarrollo muy distintas a las que se han venido aplicando por los gobiernos y por los organismos internacionales.

En particular, la teoría general conduce a políticas que modifiquen el legado histórico de la desigualdad, pues la mayoría de los países del tercer mundo se parecen a la sociedad sigma. Cambiar la distribución de los activos económicos y sociales

que ha prevalecido por décadas o centurias es todo un desafío. Construir una sociedad homogénea en términos de derechos es una forma de hacerlo. Iguales derechos frente a la ley, frente al acceso a los bienes públicos, a la ciudadanía, y a las expresiones culturales —que eliminen no tanto la multiculturalidad sino su jerarquía—. Esta política significaría aplicar un *shock* refundacional en la sociedad.

Pero, ¿quiénes serían los agentes que tienen poder e incentivos para hacerlo? Esta investigación muestra que el problema dual de convergencia se debe a fallas del mercado y de los gobiernos. Los capitalistas no tienen incentivos para llevar a cabo el *shock* refundacional. El orden social que se generaría con este cambio los beneficiaría, pero el orden social es un bien público, sujeto al problema olsoniano de la acción colectiva. Los trabajadores tienen el incentivo pero no tienen el poder, pues enfrentan restricciones severas para llevar a cabo la acción colectiva que se requeriría, además del problema olsoniano.

Los gobiernos no tienen incentivos para llevar a cabo el *shock* refundacional. Los estudios de economía usualmente culminan con un conjunto de recomendaciones de política para los gobiernos. En este estudio, no se hace tal cosa, porque esta misma investigación muestra que los gobiernos tienen también sus propios intereses y actúan motivados por esos intereses. Las acciones de los gobiernos son endógenas al proceso económico. Aún más, el establecer derechos para crear sociedades socialmente homogéneas es incompatible con los incentivos de los gobiernos. No es difícil, entonces, comprender por qué la historia de las políticas públicas es la historia de los fracasos de los gobiernos para generar desarrollo.

Las organizaciones internacionales han revelado a través de sus acciones que tampoco tienen interés en este tipo de reformas. Las reformas liberales que se implementaron a escala casi planetaria en los años ochenta y noventa no incluyeron una redistribución de activos y, por consiguiente, no han dado los resultados esperados en cuanto a desarrollo económico, tal como

lo predice la teoría general. Esta conclusión se puede colegir del *Informe del Banco Mundial del 2000/2001* (World Bank 2001). Las políticas del Fondo Monetario Internacional están muy ausentes de la cuestión de la desigualdad, tal como se puede leer en los relatos de sus experiencias concretas que hace Stiglitz (2002).

En cuanto a la política de desarrollo económico, la economía estándar sugiere que la inversión es fundamental. Lo que los países del tercer mundo necesitan para desarrollarse es invertir en capital físico y en capital humano. Señalan, así, que hay una relación entre inversión y desarrollo. Pero, a la luz de la teoría general, la inversión es apenas una causa próxima del desarrollo económico. ¿De qué dependen las inversiones? Estas dependen de las condiciones iniciales de la sociedad, incluyendo entre estas a la desigualdad inicial. Las condiciones iniciales constituyen la causa última. La teoría general no puede explicar los cambios en la desigualdad inicial porque esta es una variable exógena del proceso económico. La pregunta sobre quiénes serían los agentes que podrían lleva a cabo el *shock* refundacional es, en consecuencia, una cuestión abierta.

El libro también da lugar a un conjunto de implicaciones sobre el papel del conocimiento científico en el desarrollo económico. En comparación a la Física, la Economía estudia procesos más complejos. Por ejemplo, según la teoría general, en el proceso económico del desarrollo de los países, la historia cuenta. Por lo tanto, la ontología universalista de la Física no puede ser aplicada a la Economía como si el proceso económico fuera fundamentalmente el mismo en todas las sociedades. En la Física, consecuentemente, el progreso de la ciencia se da en el primer mundo y lo único que el tercer mundo tiene que hacer para desarrollarse es aplicarlo a la solución de sus problemas (cuestión de pura Ingeniería). Eso es lo eficiente.

En cambio, para que la Economía sea útil al desarrollo del tercer mundo y pueda generar una ingeniería social del desarrollo, la ciencia económica relevante tiene que ser creada y

desarrollada. En esta tarea, se tiene que combinar teorías generales con teorías particulares. Esta es la conclusión principal de este libro. Una metodología similar se utiliza en la Biología. Por ejemplo, para explicar el comportamiento de las plantas en los Andes se tiene que combinar los principios generales de la fotosíntesis con los de la fisiología vegetal en ambientes con variaciones importantes de temperatura durante un mismo día —y no durante las estaciones, como ocurre en las zonas templadas donde se ubican la mayoría de los países del primer mundo—. La Economía y la Biología evolucionista son ciencias históricas. El método de la ontología universalista no puede ser aplicado a ellas, como se muestra en este estudio para el caso de la Economía.

Capítulo 1
La economía como ciencia social

¿Cuáles son las reglas del conocimiento científico? La teoría del conocimiento trata esta cuestión. Esta es una teoría normativa que sirve de base a las teorías positivas (que buscan explicar la realidad). Aunque pueda parecer paradójico, una teoría positiva necesita de una teoría normativa. La teoría del conocimiento es entonces una metateoría. Y, como toda teoría, es un sistema lógico. Los filósofos de la ciencia han hecho la distinción entre lo que dice la lógica de la ciencia y lo que en realidad hacen los científicos. Esa distinción también se utiliza en este trabajo.

Debido a que una teoría es un sistema lógico general, no siempre es operacional. Este es el caso de la teoría del conocimiento. Para hacerla operacional, se necesita introducir un conjunto de supuestos auxiliares que sean lógicamente consistentes con la teoría. Así, la teoría se transforma en un modelo. El modelo es, entonces, un sistema lógico que contiene reglas prácticas para acercarse al conocimiento científico. En este trabajo, se presenta un modelo particular de la teoría desarrollada por el filósofo Karl Popper. A dicho modelo lo llamaré metodología alfa-beta. La naturaleza del conocimiento en las ciencias sociales en general y en la Economía en particular será discutida a la luz de esta metodología.

Los científicos muestran desacuerdos. El origen de esos desacuerdos es variado. En algunos casos, se deben a las diferencias en la teoría del conocimiento que utilizan o en el modelo particular de la teoría del conocimiento que han adoptado, pero también a las diferencias en la teoría científica que utilizan o en la evidencia empírica. Debido a que en este trabajo se presentará una nueva teoría económica sobre el funcionamiento del

capitalismo, parece apropiado presentar de manera explícita la metodología científica que se utilizará. De esta manera, quienes estuvieran en desacuerdo con las conclusiones del libro podrán identificar las fuentes del desacuerdo.

Por lo tanto, debe quedar en claro que este capítulo no intenta discutir los problemas de la epistemología en general. Tampoco intenta dar normas sobre lo que los científicos deben hacer. Eso sería un absolutismo inapropiado. La frase común «ciencia es lo que los científicos hacen» ignora la diferencia entre la lógica del conocimiento científico y la lógica que subyace al comportamiento de los científicos.

1.1. La ciencia social como el estudio de procesos complejos

Toda ciencia se define como un sistema de proposiciones (Popper 1959: 35). Estas proposiciones deben (a) estar formuladas ordenadamente, (b) establecer relaciones entre objetos y (c) tener un criterio de validez. El ordenamiento de las proposiciones debe ser lógico —antes que taxonómico o lexicográfico—. Debe obedecer los requerimientos de un sistema lógico, requisito que incluye un conjunto finito de proposiciones primarias o axiomas, y las reglas de inferencia para deducir una proposición de otra. La ciencia es entonces el uso de silogismos.

Las relaciones entre los objetos pueden referirse a dos tipos de objetos: mentales y fácticos. En el primer caso, las relaciones se dan entre ideas, lo que deriva en las ciencias formales. En el segundo, las relaciones se establecen entre elementos del mundo real, lo que deriva en las ciencias fácticas.

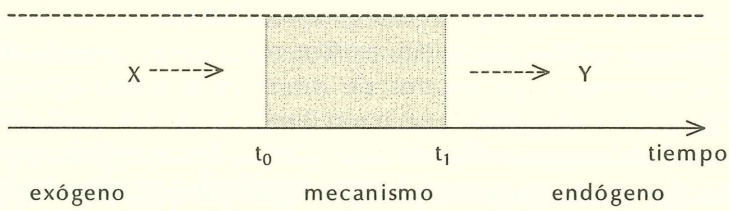
¿Cuál debería ser el criterio de verdad para alcanzar el conocimiento científico? La respuesta a esta pregunta depende del tipo de ciencia. En una ciencia formal, el que las proposiciones estén libre de contradicciones lógicas internas es condición necesaria y suficiente. En las ciencias fácticas, los criterios requieren algo adicional. Por supuesto, las proposiciones deben

ser lógicamente correctas. Pero, en este caso, ello solo es condición necesaria. Debe haber también consistencia entre teoría y realidad. Esta consistencia puede ser accidental o sistemática. La ciencia está interesada en esto último.

Las ciencias sociales pertenecen al grupo de las ciencias fácticas. En general, las ciencias sociales buscan explicar el funcionamiento de las sociedades humanas. Estas son realidades más complejas que las del mundo físico debido, en parte, a que los individuos que la conforman no son idénticos. Las ciencias sociales buscan, entonces, descubrir relaciones sistemáticas en realidades complejas. Sin embargo, no todo aspecto de la realidad social puede ser sujeto de conocimiento científico, sino únicamente aquellos fenómenos que pueden ser representados en la forma de un proceso. Para ser comprendidas, las realidades complejas deben ser reducidas a un proceso abstracto.

¿Qué es un proceso? Es una serie de actividades, de una duración determinada, encaminada a un fin y repetida periodo tras periodo. En un proceso, existen elementos que ingresan —los elementos exógenos— y elementos que salen, que son el resultado del proceso —los elementos endógenos—. La Figura 1.1 ilustra el concepto de proceso, en donde el segmento t_0-t_1 representa la duración del mismo, X representa el conjunto de elementos exógenos e Y el conjunto de elementos endógenos.

Figura 1.1
Representación diagramática de un proceso



En un proceso existe también un mecanismo subyacente por el cual los elementos exógenos afectan a los elementos

endógenos. Lo que sucede al interior del proceso —el mecanismo de interacción entre los elementos que intervienen— no puede ser observado (como se indica en la Figura 1.1 por medio del área sombreada). Si fuera observado, el interior del proceso podría ser considerado como un proceso en sí mismo, con elementos endógenos y exógenos, y un interior inobservable; y así podríamos retroceder en una regresión continua. En última instancia, debe existir algo escondido detrás de las cosas que observamos. La ciencia busca entonces descubrir estos elementos subyacentes.

Dado que un proceso se repite periodo tras periodo, las relaciones entre los elementos endógenos y los exógenos pueden ser continuamente observadas empíricamente, posibilidad que da lugar al surgimiento de relaciones sistemáticas o regularidades. La ocurrencia de regularidades en un fenómeno real es condición necesaria para que este sea sujeto de investigación científica. Un mundo caótico —en el que las regularidades estén ausentes— presenta dificultades serias para ser sujeto de investigación científica. De este mundo no se va a ocupar este trabajo.

Una ciencia social está interesada en explicar el resultado de un proceso; entonces, los elementos endógenos del proceso constituyen el objeto de estudio de la ciencia respectiva, en tanto que los elementos exógenos y los mecanismos de interacción entre estos y aquellos son los que permiten explicar los fenómenos, y son establecidos, ambos, mediante supuestos. La reducción de la realidad a un proceso abstracto es entonces realizada mediante supuestos, esto es, mediante la construcción de una teoría.

En una realidad compleja, existen muchos elementos exógenos y muchos mecanismos de interacción entre elementos endógenos y exógenos que intervienen en su funcionamiento. En este sentido, la realidad es compleja. Por consiguiente, una teoría toma en cuenta únicamente algunos mecanismos y algunos elementos exógenos, aquellos que se supone son los factores esenciales que influyen en los elementos endógenos de la realidad. El resto de elementos exógenos es ignorado en tanto

se supone que ellos son circunstanciales, anecdóticos. Esta simplificación es inevitable, por la misma razón que un mapa a escala 1:1 es inútil en geografía. Por necesidad lógica, se necesita de una teoría para transformar o reducir la realidad a un mundo abstracto y simple. Para sustituir el mundo real por otro abstracto, una teoría necesita ser un sistema de relaciones que sea lógicamente correcto, en tanto se supone que el mundo real funciona bajo esta lógica.[1]

Los supuestos de una teoría no necesitan justificación. Si tuvieran que justificarse, se necesitaría otro conjunto de supuestos, y estos a su vez necesitarían otros, y así sucesivamente; por tanto, estaríamos retrocediendo en una regresión continua. La abstracción es reduccionismo, pues la realidad se reduce a sus elementos supuestamente esenciales. Pero los supuestos se eligen como parte de un algoritmo para llegar al mejor conjunto de supuestos, es decir, a la mejor teoría. Si los supuestos iniciales fallan se les abandona y reemplaza por otros. Los supuestos de la primera teoría serán arbitrarios, pero no así los de las teorías siguientes, pues no se pueden repetir supuestos que ya se mostraron equivocados. No existe criterio a priori para demostrar que el conjunto de supuestos N es superior al conjunto M. Es, por lo tanto, de poca utilidad defender o atacar una teoría sobre la base de sus supuestos. Los supuestos no necesitan justificación sino confrontación con la realidad.

En las ciencias sociales, una teoría es un conjunto de supuestos acerca de las fuerzas y motivaciones que subyacen a los hechos observados. Una teoría no es, por tanto, directamente observable. Es un sistema lógico, libre de inconsistencias lógicas internas.

---

[1] La antigua proposición acerca de que la Tierra es un plato que reposa en el caparazón de una tortuga gigante no es un argumento lógico si no se nos dice sobre qué descansa la tortuga. Hawking (1988) cuenta la historia que alguien pretendió resolver este problema diciendo «...sobre otra tortuga». Esta proposición, por ejemplo, no puede servir de base para una teoría del universo.

¿Qué es observable (medible) y qué no lo es en el proceso abstracto? Esta es una cuestión central en el conocimiento. Los elementos endógenos de un proceso deben ser observables, por lo que son los elementos del análisis empírico. Así, estos elementos constituyen las *variables endógenas,* en el sentido de que sus variaciones son observables. Los elementos exógenos que forman parte de un proceso pueden ser separados en dos categorías: observables y no observables. Aquellos que son observables pueden definirse como las *variables exógenas.* Como se mencionó líneas arriba, el mecanismo de interacción es inobservable.

El objetivo de las ciencias sociales es establecer *relaciones de causalidad,* esto es, determinar *qué causa qué.* La causalidad es lo que hace a la ciencia útil; le otorga a las personas poder para alterar la realidad. Las relaciones de causalidad indican el efecto de los cambios en las variables exógenas sobre las variables endógenas. Por lo tanto, la causalidad se deriva lógicamente de una teoría. Pero, entonces, las relaciones de causalidad deben ser observables. Viniendo de procesos complejos, las relaciones de causalidad en las ciencias sociales se refieren a regularidades y no a relaciones anecdóticas.

Las características observables del mundo real pueden ser explicadas (la respuesta a la pregunta del «por qué») solamente sobre la base de factores no observables, esos que son subyacentes. Este es el papel de la teoría. No podría darse una explicación de un hecho observado sobre la base de otro hecho observado. Esta última relación empírica es, más bien, una relación de causalidad que solo tiene sentido lógico a la luz de una teoría.

1.2. La metodología alfa-beta

En términos de la ordenación lógica de las proposiciones, el concepto operacional de la ciencia que se utiliza en este traba-

jo se basa en la definición dada por Georgescu-Roegen (1971). Esta definición es la siguiente: *ciencia es un conjunto de proposiciones alfa y beta, tal que las proposiciones beta son derivadas lógicamente de las proposiciones alfa, y ninguna proposición alfa puede ser derivada de otra proposición alfa.* Así, los fundamentos de la ciencia vienen dados por sus proposiciones alfa.

El conjunto de proposiciones alfa constituye la teoría. Es el conjunto de supuestos o axiomas que se establece para comprender la realidad. Estos supuestos se construyen como axiomas, no necesitan justificación. Explicar la realidad (respondiendo a la pregunta del «por qué») requiere teoría o abstracción, y eso implica simplificación de la realidad. Las teorías siempre representarían versiones simplificadas de la realidad. Cómo encontrar la mejor simplificación de la realidad (el mejor conjunto de supuestos) es un problema metodológico y no una cuestión que pueda ser resuelta mediante la justificación de las razones que dieron lugar a su selección.

Las proposiciones beta se obtienen de las alfa por inferencia lógica. Las proposiciones beta muestran las relaciones lógicas entre las variables exógenas y endógenas. Por lo tanto, las proposiciones beta indican las implicancias empíricas de la teoría. Si una teoría es una buena aproximación de la realidad, se debería observar ciertas relaciones particulares en el mundo real, aquellas que predice. Aunque el conjunto de proposiciones alfa se establece de manera arbitraria, pues son los supuestos, el conjunto de proposiciones beta ya no puede ser arbitraria. Las proposiciones beta, aunque lógicamente correctas, pueden ser empíricamente falsas.

Teoría y causalidad pueden ser expresadas en correspondencia con estos dos tipos de proposiciones: la teoría corresponde a las proposiciones alfa y la causalidad a las proposiciones beta. Las relaciones de causalidad son, entonces, lógicamente derivadas de la teoría. Las proposiciones beta tienen contenido empírico, predicen relaciones empíricas particulares entre las variables exógenas y las endógenas. Los efectos —positivos o

negativos— de las variables exógenas sobre cada una de las variables endógenas son derivados de la teoría. Las relaciones causa-efecto toman la forma de proposiciones beta. En consecuencia, para cada variable endógena existirá una relación causal; así, habrá tantas proposiciones beta como variables endógenas contenga el sistema teórico.

Una proposición beta es una proposición empíricamente refutable; esto es, en principio puede ser empíricamente falsa. Una buena teoría es aquella que da lugar a proposiciones beta que, pudiendo ser empíricamente falsas, cuando son contrastadas con la realidad, resultan ser consistentes con los datos de la realidad (una buena teoría es como una persona honrada que, pudiendo delinquir, no lo hace). Si una teoría da lugar a proposiciones beta que nunca pueden fallar empíricamente, esta teoría es una tautología, inútil para el conocimiento científico.

En las ciencias sociales, la teoría es entonces un conocimiento a priori, un artificio lógico que nos permite arribar al conocimiento científico. Si no hay teoría, no puede haber conocimiento científico. Pero la teoría necesita confirmación empírica. El conocimiento a priori necesita una confirmación a posteriori. ¿Por qué? Porque los supuestos fueron establecidos arbitrariamente. Si, en esta confrontación, teoría y realidad son inconsistentes, es la teoría la que pierde (no la realidad); esto es, queda demostrado que la selección arbitraria de sus supuestos fue errónea. Aun cuando una teoría es un sistema lógicamente correcto, puede ser empíricamente falsa.

El conjunto de proposiciones alfa no puede ser sujeto de análisis empírico, pero sí lo pueden ser las proposiciones beta. Así, la confrontación empírica de una teoría puede realizarse solo de forma indirecta, a través de las proposiciones beta. Este hecho es posible porque las proposiciones beta tienen contenido empírico, predicen una relación particular y observable entre las variables endógenas y exógenas, predicen relaciones de causalidad. Son estas predicciones empíricas de la teoría las que pueden o no coincidir con la realidad. Cuando esta consis-

tencia se da, se dice que la teoría explica la realidad. El mundo abstracto de la teoría coincide con el mundo real. El mundo real se puede comprender, en lo esencial, a partir del mundo abstracto de la teoría.

A causa de su papel empírico, las proposiciones beta excluyen los elementos exógenos inobservables del proceso. ¿Por qué? Porque asumiendo cambios apropiados en los elementos exógenos inobservables *siempre* se puede hacer coincidir a la teoría con la realidad. Al introducir elementos exógenos inobservables, una teoría se vuelve inmortal. En Economía, por ejemplo, se puede «explicar» cualquier variación en los precios de mercado introduciendo cambios en las preferencias o en las expectativas de los individuos. Lo que se necesita, como teoría, es un conjunto de proposiciones del que se puedan derivar proposiciones empíricamente refutables. Las proposiciones beta son, según su construcción lógica, proposiciones empíricamente refutables. Por esta razón, las proposiciones beta incluyen únicamente variables observables. El problema epistemológico de establecer la relación de causalidad entre una teoría y su confrontación empírica queda resuelto mediante las proposiciones beta.

El proceso abstracto presentado en la Figura 1.1 se refiere únicamente a una parte de la realidad. Un proceso es siempre una representación parcial de la realidad. Por consiguiente, una ciencia social específica (Antropología, Ciencia Política, Economía o Sociología) tiene que ver con un aspecto particular de la realidad. Pero el conocimiento científico requiere unidad de conocimiento —un único mundo abstracto, una unidad ontológica—.

¿Cómo podrían todas las ciencias sociales generar un conocimiento agregado consistente, esto es, una unidad en el conocimiento? Este problema epistemológico aún no ha sido resuelto ni en las ciencias naturales. Recientemente, el biólogo Edward Wilson ha discutido el tema y ha argumentado que la unidad del conocimiento puede lograrse reduciendo todas las ciencias

fácticas a las leyes de la Física porque los organismos vivientes y las organizaciones sociales funcionan *como si* fuesen máquinas (Wilson 1998).

Una solución alternativa podría provenir de una división del trabajo coordinada entre las ciencias sociales: lo que es exógeno en determinada ciencia será endógeno en otra. Todas las ciencias fácticas serían complementarias. La unidad del conocimiento se establecería, entonces, a través de la consolidación lógica de las relaciones causales establecidas en cada ciencia. Todas las variables endógenas en el sistema de las ciencias sociales dependerían de las variables exógenas del ahora único mundo abstracto.

El problema de la unidad del conocimiento podría también surgir al interior de una misma ciencia fáctica. Como se sabe, la Física enfrenta el problema de la unidad del conocimiento. La teoría de la mecánica cuántica (el mundo de lo pequeño) no es consistente con la teoría de la relatividad (el mundo de lo grande). El desarrollo reciente de la teoría de las cuerdas intenta reconciliar esta inconsistencia (Hawking 2002). En Economía, la teoría microeconómica muestra, en muchos aspectos, inconsistencia con la teoría macroeconómica. La moderna *teoría de equilibrio general* intenta reconciliar estas inconsistencias, como se verá en este estudio.

La metodología de las ciencias fácticas se refiere al sistema lógico que sirve de guía en las decisiones para aceptar o rechazar teorías. La metodología es el conjunto de reglas para aceptar o rechazar teorías. Sobre la base de las proposiciones alfa-beta se desarrolla aquí una metodología particular.

En las ciencias sociales, dado que una teoría necesita hacer abstracciones, algunos de los elementos que influyen en las variables endógenas son dejados de lado, por lo que los datos empíricos difícilmente pueden calzar a la perfección con las proposiciones beta. Esta falta de precisión es el precio que se tiene que pagar por la simplicidad de una teoría explicativa.

Es lógicamente incorrecto esperar que una proposición beta —derivada de una teoría, de una abstracción de la realidad— sea consistente con todos los casos observados. La consistencia empírica solo puede darse con los casos generales o típicos. Así, se requiere del análisis estadístico para aceptar o rechazar una proposición beta. Encontrar una sola observación en la que la prueba falle es insuficiente para rechazar una proposición beta, en tanto el valor estadístico de una observación es nulo. Esa observación podría responder a un error, a una desviación de la norma. El test debe ser, entonces, estadístico. Es la relación entre los valores promedios de las variables lo que cuenta. En consecuencia, debe establecerse una distinción entre *error* y *falla* en el proceso de aceptación o rechazo de una proposición beta. La lógica de esta metodología es internamente consistente. Una realidad compleja que ha sido representada como un proceso abstracto, en el que existe repetición y las regularidades empíricas pueden suceder, es susceptible de análisis estadístico y, entonces, materia del conocimiento científico.

Si las proposiciones beta se ajustan bien a los datos empíricos, puede decirse que la teoría es *consistente* con la realidad. No podemos decir que la teoría es *verdadera*. Esto es así porque las mismas proposiciones beta podrían derivarse de otra teoría. No hay una relación de uno a uno entre las proposiciones alfa y las proposiciones beta. Por tanto, la aseveración acerca de que una teoría «explica» la realidad tiene un significado muy preciso: sus proposiciones beta son consistentes con los datos empíricos. Si los datos empíricos no coinciden con las proposiciones beta, la teoría es simplemente falsa. En este caso, se formula una nueva teoría y el algoritmo continúa.

Esta regla de iteración, basada en la acción conjunta de las proposiciones alfa y beta, será denominada la *metodología alfa-beta*. La interacción continua entre la teoría y los datos empíricos es la piedra angular de esta metodología.

La metodología alfa-beta está representada en la Figura 1.2. A partir de un conjunto de proposiciones teóricas α_1, se obtiene

lógicamente un conjunto de proposiciones empíricamente observables β_1 (indicado mediante la flecha doble), el que debe ser sujeto de análisis estadístico (indicado mediante la flecha simple). La flecha doble indica el procedimiento lógico; la flecha simple señala el procedimiento operacional, la actividad a realizar. El análisis estadístico (señalado mediante el símbolo $\approx$) implica la búsqueda de una conformidad estadística entre las proposiciones beta y el conjunto de relaciones, asociaciones o correlaciones empíricas entre las variables endógenas y exógenas, que denominaremos b. Si se muestra que *estadísticamente* $\beta_1 = b$, entonces α_1 es consistente con la realidad. Si $\beta_1 \neq \beta$, entonces α_1 es falso y debe desarrollarse una nueva teoría α_2, y se continúa así con el algoritmo.

Figura 1.2
La metodología alfa-beta

$\alpha_1 \Rightarrow \beta_1 \rightarrow [\beta_1 \approx b]$

Si $\beta_1 = b$, α_1 es consistente con b y explica la realidad
Si $\beta_1 \neq b$, α_1 es falso y no explica la realidad. Entonces,

$\alpha_2 \Rightarrow \beta_2 \rightarrow [\beta_2 \approx b]$

Si ... (el algoritmo continúa)

El conjunto de las relaciones o asociaciones empíricas (indicadas por la letra b en la Figura 1.2) se calcula a partir de una base de datos a la cual se le aplica una teoría estadística de inferencia. Por lo tanto, el conjunto de asociaciones empíricas obtenidas cambiará cuando se cuente con una nueva o mayor base de datos o con mejores teorías estadísticas. Innovaciones en los instrumentos de medición también modificarán la base de datos. Estos casos de modificaciones darán lugar a nuevas pruebas estadísticas sobre la validez empírica de la teoría.

Dado el tamaño del vector b, una teoría resultaría falsa si una proposición beta es falsa. Así, una teoría es empíricamente consistente con la realidad si y solo si ninguna de sus proposiciones beta es falsa.

Cuando se confrontan empíricamente varias teorías a la vez, dada una base de datos, algunas de ellas resultarán falsas y otras resultarán consistentes. Con nueva información, algunas de las teorías del último grupo serán falsas y otras consistentes. Así sucesivamente. Aquellas que sobreviven a todos los procesos de confrontación serán las teorías aceptadas. Y reinarán hasta que una nueva información disponible conduzca a la aparición de inconsistencias.

Un ejemplo sencillo puede ilustrar la metodología alfa-beta. Suponga una figura F que no es observable como figura por el investigador. Considere la siguiente teoría 1 que propone un investigador: «la figura F es un cuadrado». De serlo, entonces sus dos diagonales deberán ser iguales (una proposición beta). Si las dos diagonales son estadísticamente iguales, la teoría es consistente con la realidad. Sin embargo, el investigador no puede afirmar que la figura F es un cuadrado, dado que esta evidencia es también consistente con la teoría 2 que establece lo siguiente: «la figura F es un rectángulo». Las teorías 1 y 2 siguen siendo consistentes con esta información.

Suponga que la información inicial hubiese sido otra. El dato muestra que las diagonales no son iguales. En este caso, las teorías 1 y 2 son inequívocamente falsas. En este caso, el investigador tendría que desarrollar otra teoría, digamos la teoría 3, que establece que «la figura F es un rombo». Esta teoría es ciertamente consistente con esta información.

Realmente, no hay una secuencia particular implícita en la metodología alfa-beta. El programa de investigación podría empezar con observaciones empíricas, con asociaciones estadísticas, o con la construcción de una teoría (un mundo puramente abstracto). Lo que es importante es seguir la lógica del método, es decir, el algoritmo indicado.

Una proposición beta es una proposición sintética: puede ser o verdadera o falsa, en términos estadísticos. Una buena teoría es fácilmente identificable: aun cuando sus predicciones podrían haber resultado descalificadas por la observación empírica, estas han sido confirmadas. Una teoría necesita ser empíricamente refutable; en principio, una teoría debe ser mortal. Si una proposición no es susceptible de generar proposiciones refutables, no puede convertirse en teoría. Se necesita falsar una teoría, es decir, contrastar sus proposiciones beta con los datos de la realidad de forma que exista la posibilidad de refutarla.

Este es el principio de falsación *(principle of falseability)* que estableció Popper —principio que constituye la línea demarcatoria entre ciencia y no ciencia— (Popper 1959). «Los hombres mueren cuando los dioses lo determinan» es un ejemplo de una proposición que no genera predicciones refutables, pues la voluntad de los dioses no es observable. Este tipo de proposición es una tautología, inservible para la investigación científica.[2]

La metodología alfa-beta implica que una teoría está sujeta a continuas confrontaciones con la realidad. La buena teoría es aquella que ha sobrevivido a numerosos intentos de refutación. Aceptar teorías es siempre una decisión provisional, pero su rechazo es definitivo. Las proposiciones beta nos dicen lo que hay que observar para refutar una teoría (no solo para confirmarla). Este procedimiento elimina cualquier posibilidad de estratagemas inmunizadoras que salvaguarden a la teoría de la refutación empírica.

A partir de la metodología alfa-beta, debe quedar en claro que los datos por sí mismos no pueden explicar fenómenos rea-

[2] Los curanderos son particularmente buenos en el uso de proposiciones no refutables en su trato con personas enfermas: «Si tienes fe, sanarás». Cuando una de las personas se lamenta de no estar mejorando, el curandero responderá «No tuviste fe en mi medicina». Los curanderos nunca fallan.

les; estos no pueden producir relaciones de causalidad. Producirán correlación estadística, pero esto no es causalidad. No existe una ruta lógica desde la asociación o correlación empírica hacia la causalidad. Esta última requiere una teoría subyacente, dado que las variables exógenas y las variables endógenas vienen de la teoría. No puede existir entonces un «test de causalidad» que, a partir de la información empírica, establezca que las variables A son la causa de los cambios en las variables Z.[3]

Puede ocurrir que las proposiciones beta no puedan ser derivadas directamente de las proposiciones alfa debido a que estas son proposiciones muy generales. En este caso, se requiere transformar las proposiciones alfa en un conjunto de proposiciones operacionales que permita derivar de él proposiciones beta. Este conjunto de proposiciones se llama *modelo* de la teoría. Un modelo teórico es un conjunto de supuestos que incluye las proposiciones alfa y un subconjunto de *supuestos auxiliares consistentes.* La consistencia se refiere a que los supuestos auxiliares no pueden contradecir los supuestos básicos de la teoría. Una teoría es entonces una familia de modelos, donde el conjunto de proposiciones alfa constituye el elemento común de toda la familia.

Para una teoría α, un modelo α_j de esta teoría puede representarse de la siguiente forma: $\alpha_j = (\alpha, a_j)$, donde a_j es un conjunto de supuestos auxiliares que son consistentes con α y donde $j=1, 2, ..., n$. Un modelo es también un sistema lógico, libre de inconsistencias lógicas internas. En consecuencia, los supuestos auxiliares no pueden ser arbitrarios.

Si en una ciencia social dada, el conjunto de proposiciones alfa no es directamente operacional, se necesita introducir un

[3] El así conocido «test de causalidad de Granger» recibe un nombre inapropiado. Debería llamarse «test de secuencialidad de Granger», dado que prueba si cambios en una variable consistentemente preceden los cambios en alguna otra, donde el número de periodos de rezagos de la prueba estadística es arbitrario.

conjunto de supuestos auxiliares y consistentes adicional para hacer la teoría refutable. La lógica de la construcción de los modelos es falsar la teoría.

En la Figura 1.2, considere que α_1 y α_2 sean dos modelos de la teoría α. El algoritmo de falsación es aplicado entonces a estos modelos. Si el primer modelo no pasa la prueba, el segundo ingresa a la prueba, y así sucesivamente. Para que una teoría sea falsa, todos los modelos de la familia deben fallar. El algoritmo requiere, por tanto, que el número de modelos sea finito; esto es, se requiere que el número de supuestos auxiliares sea limitado. Si todos los modelos fallan, la teoría es empíricamente falsa y es necesario construir una nueva teoría.

De la metodología alfa-beta se sigue lógicamente que el conocimiento científico es siempre provisional. Una teoría exitosa se convertirá en lo que Kuhn denomina el «paradigma». Pero este paradigma tendrá una vida temporal. Reinará hasta que aparezcan nuevas evidencias empíricas que descalifiquen sus predicciones.

La metodología alfa-beta es consistente con la teoría del conocimiento de Popper. Esta teoría sigue su principio de demarcación: cualquier proposición que en principio no puede ser refutable está fuera del dominio de la ciencia (Popper 1959). La proposición beta incorpora el principio de *falsabilidad* como criterio para aceptar o rechazar una teoría. La metodología alfa-beta es un modelo de la teoría de Popper. También incorpora la metodología del joven Feyerabend, pues el algoritmo que usa se basa en el uso del *pluralismo teórico*, y los modelos juegan el papel de cinturones de protección (finitas) de las teorías (Brown 2001).

De acuerdo con la metodología alfa-beta, es suficiente la falla estadística de una proposición beta para rechazar una teoría que pretende explicar cierta realidad (una vez que se demuestra que las dos diagonales son diferentes, no hay necesidad de pruebas adicionales para rechazar la teoría que establecía que la figura F es un cuadrado). En el caso de la existencia de

modelos, la teoría es falsa si las proposiciones beta que se refieren a una sola variable endógena en todos los modelos son falsas.

Como señala Blaug, algunos epistemólogos han exagerado el así llamado problema Duhem-Quine: la imposibilidad de poner a prueba, conclusivamente, cualquier teoría científica. El argumento es que la introducción de supuestos auxiliares y condiciones *ceteris paribus* hacen posible que cualquier teoría pueda librarse de la refutación. El método de Popper intenta resolver este problema: «No es la falsabilidad lo que distingue la ciencia de la no ciencia en Popper; lo que demarca a la ciencia de la no ciencia es la falsabilidad unida a reglas metodológicas para evitar [...] estratagemas inmunizadoras» (Blaug 1992: 19). En la metodología alfa-beta, estas reglas incluyen la introducción de supuestos auxiliares consistentes. Estos supuestos, por tanto, no pueden disminuir el grado de falsabilidad de una teoría, sino que por el contrario lo incrementan.

La condición *ceteris paribus* está claramente definida en la metodología alfa-beta. Solamente las variables exógenas (observables) son consideradas variables explicativas. Los elementos exógenos inobservables no son considerados en el proceso de falsación, pues harían inmortal a la teoría. Por lo tanto, «todo lo demás constante» en un proceso de falsación solamente podrá significar que algunas variables exógenas son mantenidas fijas de acuerdo al diseño experimental.

Para ser claros, la metodología alfa-beta no usa la metodología del *verificacionismo,* que se basa en la selección de los datos empíricos con el fin de confirmar la teoría. En este caso, el investigador busca datos para confirmar la teoría, no para refutarla.[4] También descarta el *interpretacionismo* —llamado tam-

[4] El físico Richard Feynman se refiere a este problema como falta de integridad científica. Alguna vez le dijo a sus estudiantes: «Si están llevando a cabo un experimento, deben reportar todo lo que crean que lo invalida —no solo lo que creen que es favorable a la teoría—. [...] La idea es tratar de dar *toda* la informa-

bién *hermenéutica*—, puesto que este hace a las teorías inmortales. La metodología alfa-beta también excluye el uso del *aplicacionismo* —trabajo empírico dirigido a explicar un fenómeno utilizando la teoría—, ya que este parte del supuesto de que la teoría explica la realidad en cuestión. También descarta el *empirismo,* en tanto las asociaciones estadísticas per se (coeficientes de correlación y análisis de regresión multivariada, y cualquier otra técnica sofisticada de asociación) no pueden generar causalidad; esta solo se deriva de un sistema lógico, de una teoría.

El posmodernismo ha tenido alguna influencia en la metodología —sobre todo en la metodología de la Economía— a través del uso de la *retórica.* Bajo esta metodología, «todo vale». Una teoría puede ser defendida sobre la base de la persuasión, esto es, mediante la utilización de algunas piezas de evidencia, casos históricos, sentido común y, más aún, la teoría misma. La metodología alfa-beta no usa la retórica.

En cuanto a las pruebas estadísticas, la metodología alfa-beta requiere que estas se basen en pruebas no paramétricas. El análisis estadístico convencional está basado en una teoría estadística paramétrica, que —como cualquier teoría— se basa en supuestos. Dichos supuestos, entonces, pueden ser empíricamente falsos. Sería necesario probar la validez empírica de esta teoría estadística. Esta necesidad, a su vez, requeriría otras pruebas estadísticas, hecho que nos llevaría hacia atrás en una regresión continua. Probar estadísticamente la validez de las pruebas estadísticas constituye una imposibilidad lógica. Pero también constituye una imposibilidad práctica, porque la Estadística es una ciencia formal.

---

ción que ayude a otros a juzgar el valor de su contribución; no solamente la información que encamina el juicio en una dirección particular o en otra. [...] De lo que estoy hablando no es únicamente una cuestión de no ser deshonesto, es más una cuestión de integridad científica, lo que está en otro nivel» (Feynman 1985: 341).

El rechazo de una teoría a partir del uso de pruebas estadísticas paramétricas puede ser el resultado de fallas en los supuestos de la teoría estadística subyacente y no en los supuestos de la teoría bajo observación. La falsación requiere el uso de métodos estadísticos no paramétricos para evitar este problema de tener supuestos fuera de la teoría.

En suma, la epistemología vista como teoría del conocimiento nos brinda la posibilidad de contar con una lógica del conocimiento científico. Para ser operacional, esta teoría debe ser transformada en un modelo. A partir de este se obtiene un conjunto de reglas para aceptar o rechazar teorías en las ciencias sociales. La metodología alfa-beta es un modelo particular de la teoría popperiana del conocimiento. Establece un conjunto de reglas útiles a ser aplicadas cuando se realiza una investigación científica. Los científicos sociales pueden decidir seguir una metodología que se base en una teoría del conocimiento o pueden decidir utilizar cualquier otra metodología. Existe, entonces, una distinción entre las reglas de la metodología científica y la lógica que guía las acciones de los científicos.

La historia de la ciencia puede decirnos lo que los científicos han estado haciendo. Los trabajos de Kuhn (1970) y Galison (1997) intentan describirnos y darnos explicaciones sobre el comportamiento general de los científicos. Los cambios en los paradigmas (teoría-metodología dominante en un periodo dado) provienen, según Kuhn, de procesos sociales y políticos, mientras que, para Galisson, provienen de innovaciones en la tecnología de las mediciones empíricas. Así, la aparición del primer telescopio significó una revolución en la Física.

Se puede decir que el falsacionismo es comúnmente aplicado en la Física, al menos esto es lo que se puede deducir del famoso libro de Hawking (1988). En la Física, al parecer, las teorías son empíricamente refutadas y nuevas teorías reemplazan a las anteriores. Como resultado, Ptolomeo, Copérnico, Galileo, Newton y hasta Einstein han ido muriendo teóricamente. No en el sentido de que sus teorías fuesen totalmente erróneas

—pues algunas de las proposiciones beta de estas teorías son consistentes con la realidad—, sino en el sentido de que las nuevas teorías tienen un mayor número de proposiciones beta empíricamente consistentes, incluyendo las antiguas. Las nuevas teorías son superiores.

Pero el algoritmo continúa hasta el presente. Los físicos modernos continúan en su intento de refutar teorías, como es el caso de Prigogine (1997), quien sostiene que la Física actual tiene una predicción que es inconsistente con la realidad: predice que los fenómenos físicos son reversibles con respecto al tiempo (una taza rota puede volver sana a la mesa, un huevo frito se puede volver huevo nuevo, como si se tratara de pasar escenas de una película retrocediendo la cinta). Propone, entonces, una nueva teoría que logre predecir la irreversibilidad del tiempo.

Si bien la Física utiliza la falsación popperiana, no utiliza la metodología alfa-beta. La falsación no la hacen a través de las proposiciones beta. La razón es simple: en el proceso físico no parece existir el conjunto de variables exógenas. Por lo tanto, las predicciones empíricas de las teorías en la Física son predicciones incondicionales, es decir, son *leyes,* sean deterministas o estadísticas, pero relaciones incondicionales. Así, se ha llegado a proponer recientemente la *teoría del todo* (cf. Hawking 2002). Las proposiciones beta son, por contraste, predicciones condicionales.

En Biología, la metodología de la falsación es aparentemente menos usada. Mayr (1991) muestra que teóricamente Darwin todavía vive. Ello ocurre a pesar de lo que el biólogo Wilson (1998) nos confiesa:

> Las teorías científicas [...] son construidas específicamente para ser dejadas de lado si se demuestra que son equivocadas, y si es así, cuanto más rápido ocurra, mejor será. 'Cometan sus errores rápidamente' es una regla para la práctica científica. Concedo que con frecuencia los científicos se enamoran de sus construcciones. Lo sé, me ha

> pasado. Ellos invierten toda una vida tratando vanamente de apuntalarlas [...]. En ese caso —como el economista Paul Samuelson alguna vez bromeaba— de funeral en funeral, la teoría avanza. (1998: 52)

La metodología de la Economía se discutirá en detalle más adelante. Pero es claro que, en Economía, Smith, Ricardo, Marx, Walras, Keynes y Lucas mantienen su vigencia teórica, a pesar de que sus teorías son contradictorias entre sí.

La teoría de Kuhn predice que, en cada ciencia, en un periodo dado, habrá una teoría dominante —el paradigma—. Su teoría busca explicar los cambios en esos paradigmas. Las observaciones mencionadas muestran que, por el contrario, en la mayoría de las ciencias fuera de la Física la predicción empírica de la teoría de Kuhn —el que varios paradigmas no puedan coexistir— parece no sostenerse.

¿Por qué hacen los científicos lo que hacen? Esta es una pregunta para ser contestada teóricamente. ¿Por qué coexisten «investigadores kuhnianos» (quienes operan bajo un paradigma dado) e «investigadores popperianos» (quienes buscan refutar paradigmas)? El comportamiento de los científicos puede, consecuentemente, ser sujeto de investigación científica también. Los historiadores de la ciencia han debatido este asunto, pero sin haber aplicado un método científico en su investigación. Por ejemplo, no hay en el libro de Kuhn una teoría y su correspondiente falsación sobre el comportamiento de los científicos.

Hay todavía una cuestión adicional: ¿qué hacen los no científicos para obtener conocimiento? Considérese la proposición de Popper: «Todos los organismos vivientes están hechos para resolver problemas, aquellos problemas que surgen con la vida misma» (Popper 1993). ¿Cómo hacen estos organismos para resolver sus problemas? «Todos los organismos vivientes usan teorías y la corroboración empírica para solucionar sus problemas», añade Popper. Los que usan teorías falsas pagan el costo de su extinción. Esta puede ser la base de la selección natural.

La proposición de Popper implica que una ameba, un hombre de la calle y un científico usan teorías para resolver problemas. Lo que distingue al científico es que su metodología para eliminar teorías falsas es más desarrollada: es elaborada conscientemente, no está basada en instintos.[5] Por esto, la ciencia necesita una metodología explícita, impersonal, lógica y perfectible. La metodología alfa-beta es una forma particular de presentar dicha metodología. Como resultado, se puede establecer una diferencia entre el conocimiento científico y el conocimiento simple, aquel de los organismos vivientes.

Ciertamente, la proposición de Popper es refutable; es una buena teoría. Si esta teoría fuese consistente con los datos, es decir, si realmente todos los organismos vivientes construyen sus teorías para sobrevivir y actuar, podría afirmarse que «no hay nada más práctico que una buena teoría».

Si los organismos vivientes actúan de acuerdo a sus teorías, la tarea de los científicos es descubrir estas teorías. La metodología alfa-beta sirve para esta tarea. Por iteraciones, el científico es capaz de construir un mundo abstracto lo suficientemente parecido a aquel en el que opera y se adapta el organismo viviente. La Física no enfrenta este problema. Los átomos no pueden aprender a comportarse de modo distinto. Esta es tal vez una de las razones que hace de la Física la ciencia más desarrollada. Pero es una ciencia distinta, y una más sencilla.[6] Por otro lado, en

---

[5] La otra diferencia es que mientras las teorías de la ameba se construyen en el organismo de la ameba, el científico puede formular sus teorías mediante el lenguaje. Puede de este modo trasladar sus teorías fuera de su organismo. Así, el científico puede estudiar a todos los organismos, inclusive a su propia especie, como en las ciencias sociales. En este caso, el científico tendrá que ser el etólogo de su propia especie.

[6] Luis Alvarez, un físico americano, Premio Nobel de Física en 1968, dijo, en el discurso dirigido a los universitarios la tarde de su premiación, «la Física es la más simple de todas las ciencias. Solo parece difícil porque los físicos hablan entre ellos en un lenguaje que la mayoría de las personas no entiende —el lenguaje de las matemáticas—. Lo que hace simple a la Física es que cuando

un reciente libro que lleva por título *The End of Certainty,* Prigogine (1997) muestra que el periodo de una ciencia simple, entendida bajo condiciones de certidumbre, ha llegado a su fin. Expone, luego, una nueva teoría que supone incertidumbre y en la que las probabilidades juegan un papel central en las leyes de la naturaleza. Así, la Física parece devenir en una ciencia más compleja, más cercana a las ciencias sociales.

1.3. La metodología alfa-beta en la Economía

La Economía estudia un proceso particular: el proceso económico. Este es definido como el proceso de producción de bienes y su distribución entre los distintos grupos sociales que forman las sociedades humanas. La Economía busca establecer las variables exógenas que explican el resultado de la producción y distribución en sociedades humanas particulares. Una teoría económica debe incluir entonces un conjunto de supuestos sobre las fuerzas y motivaciones que subyacen a los hechos observados en la producción y distribución en estas sociedades.

La metodología alfa-beta se propone tener validez en todas las ciencias sociales. Pero podemos distinguir cinco particularidades de esta metodología en el caso de la Economía.

En primer lugar, la ciencia económica supone que la naturaleza del problema económico es similar en todas las sociedades. Todas enfrentan el mismo problema de escasez, esto es, los bienes deben ser producidos, pero su producción es costosa

---

ocurre un cambio en un sistema que proviene del exterior, como por ejemplo inyectarle calor, podemos predecir fácilmente que todo se elevará de temperatura. [...] Pero en el caso de un sistema infinitamente más complicado, como la población de un país en desarrollo como la India, nadie puede hasta ahora decir qué cosa hay que inyectarle para mejorar las condiciones existentes» (The Nobel Foundation 1969, «1968 Nobel Prize for Physics – Luis Alvarez», LBL Library Home page: http://www.library.lbl.gov).

en términos de recursos escasos. Mientras que los recursos son limitados en cantidad e imponen límites a la oferta de bienes, las necesidades humanas para estos bienes son ilimitadas. Este es el problema económico. Para resolverlo, se supone que cada sociedad establece las instituciones –reglas y organizaciones (North 1990)– bajo las cuales debe operar el proceso económico.

También la Economía supone que los agentes económicos manejan una lógica que subyace a su comportamiento, es decir, tienen una racionalidad económica. Esta lógica no es la lógica general, no es la metateoría que se utiliza en las matemáticas o en la lógica del conocimiento científico (presentado en la sección anterior). Es una lógica que siguen los actores sociales, un supuesto sobre las motivaciones que guían sus acciones y que los lleva a actuar de manera consistente y haciendo evaluaciones entre fines y medios. Se supone que los agentes económicos actúan no de manera emocional, sino de manera racional, que ellos calculan las consecuencias de sus decisiones.

Dadas estas peculiaridades, la ciencia económica debe construir un proceso económico abstracto. La Economía contiene dos tipos de proposiciones alfa: universales y específicas a cada tipo de sociedad. Las proposiciones alfa universales se utilizan en la construcción de cualquier sociedad abstracta.

Las proposiciones universales serán denominadas postulados, para indicar que ellas constituyen el núcleo de la Economía.[7] Estos postulados son:

α_0 *(1). El postulado de la función de producción.* Para la producción de bienes, las sociedades están dotadas de un

[7] El término *axioma* es demasiado rígido para ser usado cuando se espera que las proposiciones jueguen un rol de instrumentos en la construcción de teorías bajo iteración. El término *postulado* parece ser mejor, más flexible. *Supuesto* es un término aún más flexible y será usado en este libro cuando se trate de teorías económicas específicas.

nivel de conocimiento tecnológico y de cantidades de recursos. Los bienes no se producen de la nada.

α_0 (2). *El postulado de escasez.* Todas las sociedades buscan resolver el problema económico: los recursos para la producción de bienes son limitados, en tanto que las necesidades humanas por estos bienes son ilimitadas. El resultado del proceso económico es la producción y distribución de bienes.

α_0 (3). *El postulado institucional.* Para resolver el problema económico, todas las sociedades buscan establecer un conjunto de reglas y organizaciones para llevar a cabo el proceso económico.

α_0 (4). *El postulado de racionalidad.* El comportamiento individual está guiado por una racionalidad económica que es consistente con el contexto institucional de la sociedad en la cual el individuo vive.

El resto de proposiciones alfa serán propias de cada sociedad. Una sociedad abstracta k puede ser representada por la teoría j de la siguiente manera:

$$E_{kj} = (\alpha_0, \alpha_{kj}, \beta_{kj})$$

En este vector, el primer conjunto alfa es supuesto como universalmente válido, mientras que el segundo es específico a la sociedad k. El primer conjunto implica que, en cualquier sociedad, la producción y distribución de los bienes dependen de la dotación de recursos, de la tecnología y de las instituciones. El segundo conjunto incluye supuestos específicos a la sociedad k, de forma tal que da lugar al surgimiento de las proposiciones beta, esto es, proposiciones con contenido empírico. El conjunto beta es también propio de la sociedad k.

En segundo lugar, una teoría económica es una familia de modelos. Esto es así dado que las proposiciones alfa no son lo suficientemente precisas como para derivar proposiciones beta.

Una teoría económica puede no ser operacional debido a que contiene proposiciones generales, que se encuentran en el dominio de las intuiciones. Se requiere introducir, entonces, un conjunto de supuestos auxiliares consistentes. De esta manera, se logra transformar la teoría en proposiciones operacionales, procedimiento este que da lugar a los diferentes modelos de la teoría.

En un modelo, las proposiciones alfa se combinan con supuestos auxiliares para formar proposiciones empíricamente refutables, esto es, proposiciones beta. Un modelo nos ayuda a comprobar la consistencia lógica de la teoría. Sin embargo, el principio de falsación enfrenta ahora un problema, puesto que lo que se pone a la prueba empírica es un modelo de la teoría, no la teoría misma. El conjunto de modelos es el cinturón protector de la teoría. La refutación de una teoría implica falsar todos sus posibles modelos. Ciertamente, el principio de falsación requiere que el número de modelos en una teoría sea finito.

En el lenguaje de las matemáticas, los supuestos de un modelo teórico se representan en las *ecuaciones estructurales* de un sistema de ecuaciones. La *forma reducida* de ese sistema representa las proposiciones beta del modelo.

En tercer lugar, la derivación de las proposiciones beta en Economía está basada en el supuesto de que el estado de equilibrio exista en el proceso económico. Se dice que una situación de producción y distribución está en equilibrio si, dadas las restricciones que enfrentan las personas, ninguna de ellas tiene poder para alterar tal situación; más aún, si esto fuera posible, nadie querría hacerlo. En el proceso económico, la noción de equilibrio se refiere a un sistema estático o a un sistema dinámico. En el primer caso, el modelo supone que las interrelaciones entre las variables endógenas son contemporáneas, mientras que en el sistema dinámico estas interrelaciones son intertemporales.

Bajo un sistema estático, el concepto de equilibro implica la repetición de los mismos valores de las variables endógenas,

periodo tras periodo, mientras los valores de las variables exógenas permanezcan constantes. En un sistema dinámico, en cambio, el equilibrio implica una determinada trayectoria temporal de los valores de las variables endógenas, mientras los valores de las variables exógenas permanezcan constantes. En un sistema dinámico, el solo paso del tiempo genera movimiento en los valores de las variables endógenas, hecho que no sucede en un sistema estático. Se puede suponer que el proceso económico toma la forma ya sea de un sistema estático o de un sistema dinámico. En consecuencia, la representación de un proceso, tal como el mostrado en la Figura 1.1, puede ser entendido como un proceso estático o como un proceso dinámico.

El equilibrio en un sistema estático puede ser estable o inestable. Es estable cuando el valor de la variable endógena regresa espontáneamente a su posición de equilibrio cada vez que se encuentre fuera de este (el ejemplo intuitivo es la posición de una bola en el fondo de una taza); de otro modo, el equilibrio es inestable (el ejemplo es la posición de la bola en la cúspide de una taza volteada).

En un sistema dinámico, dadas las relaciones intertemporales entre las variables endógenas y las condiciones iniciales, los valores de equilibrio de las variables endógenas de hoy determinan sus valores de equilibrio de mañana y estos los del periodo siguiente, y así sucesivamente. En el equilibrio dinámico, entonces, las variables endógenas describen una trayectoria en el tiempo. Las variables endógenas se mueven a pesar de que las variables exógenas permanecen fijas.

Bajo ciertas condiciones, un sistema estático puede ser tranformado en uno dinámico. Si el sistema estático tiene un equilibrio único y estable en cada periodo, se le puede transformar en un sistema dinámico introduciendo relaciones intertemporales entre algunas variables endógenas, posibilidad que implica establecer condiciones iniciales en esas variables. En este caso, el equilibrio dinámico no es sino una secuencia de situaciones de equilibrio estático (Hicks 1985).

Si las condiciones de equilibrio en un sistema estático son satisfechas y si este equilibrio es estable y único, el método de la estática comparativa puede ser usado para derivar las proposiciones beta; de modo similar, en el caso de un sistema dinámico, el método de dinámica comparativa puede ser usado para tal fin (Samuelson 1947). Cambios en las variables exógenas trasladarán la situación de equilibrio a una nueva posición en el sistema estático y trasladarán la trayectoria de equilibrio a una nueva curva en el sistema dinámico. De este modo, comparando las situaciones de equilibrio, uno puede determinar el efecto de una variable exógena sobre las endógenas, es decir, se puede derivar las proposiciones beta. En ausencia del concepto de equilibrio, la causalidad no podría establecerse.

Dado que la Economía busca explicar la producción y la distribución en sociedades humanas, la noción de equilibrio debe aplicarse a todo el sistema económico. La Economía debe establecer las condiciones teóricas para este equilibrio. Esto puede hacerse usando el método del equilibrio general. Sin embargo, partes de la Economía pueden estudiarse por separado, estableciendo las condiciones teóricas de equilibrio. Esta metodología se define como el método del equilibrio parcial. Aun el comportamiento de individuos que conforman un grupo social particular puede ser sujeto de análisis económico, estableciéndose las correspondientes condiciones teóricas de equilibrio. Esta metodología se define como el método del equilibrio microeconómico.

¿Pueden ser estos métodos aplicados independientemente? No. El principio de unidad de conocimiento es un requerimiento no solo para la agregación de las ciencias empíricas sino, también, para cada ciencia empírica. Así, las tres formas de equilibrio en Economía deben ser lógicamente consistentes entre sí. Como sistema lógico, entonces, una teoría económica solo puede expresarse en términos de equilibrio general.

En cuarto lugar, la evaluación empírica de una teoría económica —la consistencia entre las proposiciones beta y los datos

empíricos— requiere del uso de pruebas estadísticas. La experimentación controlada es rara vez posible en Economía. Por el contrario, el análisis estadístico debe realizarse sobre información generada en experimentos naturales, no controlados. Así, la falsación implica el uso de una muestra grande de observaciones. Una única observación es inútil.

La prueba estadística tiene que basarse en un test no-paramétrico, de modo de evitar el problema del doble test: el de la teoría misma y el de la teoría de la inferencia estadística utilizada. El test paramétrico hace supuestos (la teoría estadística) sobre la relación entre los datos obtenidos de una muestra y los datos de la población. Esta teoría estadística tendría que ser sometida a la prueba empírica, para lo cual se requeriría otro test que necesitaría hacer supuestos sobre la inferencia, y así sucesivamente iríamos en continua regresión. Pero aún más, la Estadística es una ciencia formal, de modo que sus proposiciones no son empíricamente falsables.

Recientemente se ha desarrollado la *economía experimental*; en ella, los agentes económicos son sometidos a juegos grupales con el fin de simular comportamientos. Si se utilizara esta información para falsar teorías se llegaría al mismo problema del doble test. Se tendrían que hacer supuestos (¿qué teoría?) sobre la relación entre los datos obtenidos de una muestra en un mundo artificial y los datos de la población del mundo real.

Si el análisis estadístico muestra consistencia entre las proposiciones beta y la información empírica, puede decirse que la economía real funciona como si fuera el sistema abstracto que se construyó con la teoría. De otro modo, se concluirá que el modelo falla. Si todos los modelos lo hacen, entonces la teoría falla. La Economía es, pues, una ciencia

¿En qué sentido es la Economía una ciencia social? En un doble sentido: no solamente porque se refiere al estudio de las sociedades sino, también, porque sus proposiciones teóricas se refieren a agregados de individuos, no a individuos. La razón es

que las teorías económicas son abstracciones de sociedades y personas reales. De modo que si la teoría yerra al explicar una única sociedad, la teoría no falla. Si la teoría yerra al explicar el comportamiento de un único individuo de una sociedad dada, tampoco falla. La principal tarea de una teoría económica es la explicación del comportamiento agregado, bien sea de sociedades, bien sea de individuos. Debido al uso de la abstracción en la generación de una teoría, las pruebas empíricas deben ser, por eso, estadísticas.

Esta propiedad de la Economía no plantea mayores problemas cuando muchos individuos participan en un proceso económico. El comportamiento general de este grupo será entonces factible de ser sometido a la prueba estadística. Si un individuo singular es un agente importante de todo el proceso económico (el gobierno de una sociedad o el monopolista en una industria de una economía de mercado), la Economía puede no ser capaz de explicar este comportamiento particular. Para poder evaluar estadísticamente una teoría de este tipo de comportamiento, tendrían que observarse muchos gobiernos y muchos monopolistas. Con todo, existirían gobiernos o monopolistas cuyo comportamiento no fuese consistente con la teoría.

El mismo problema se aplica al análisis de las sociedades. Una teoría económica puede explicar el comportamiento general de un grupo de sociedades, pero no de cada sociedad. Una teoría económica puede explicar el comportamiento de un grupo de individuos, pero no de cada individuo. La observación de que la persona X fuma y no tiene cáncer, no refuta la teoría de que la nicotina produce cáncer —teoría que se supone válida en términos generales—. La observación de que una expansión en la oferta monetaria en el país Y no generó inflación, no refuta la teoría de que mayor cantidad de dinero en una economía genera inflación —teoría que también se supone válida en términos generales—. En la Economía (y en las demás ciencias sociales), es posible entonces encontrarse con sociedades sin teoría, o individuos o agentes económicos sin teoría.

Puede ocurrir, asimismo, que exista más de una teoría que explique una realidad social. La relación entre teoría y realidad no necesariamente es de uno a uno. Como sabemos, este es un problema clásico en la Física. La luz, considerada una onda, actúa en algunos casos como partícula (el planeta Tierra, considerado esférico, puede algunas veces considerarse plano, sobre todo para distancias cortas). El aspecto o faceta de la luz —o de la Tierra— que se observe variará con las circunstancias. En Economía, puede existir más de una teoría para una realidad social.

¿Se comporta una sociedad específica o un agente económico específico como la teoría predice? Esto es otra cuestión empírica. Aquí la prueba empírica es sobre la sociedad o el agente, no sobre la teoría. No obstante, el test debe ser estadístico. Se necesitan muestras grandes de observaciones del comportamiento de la sociedad o del agente. Si, en este caso, la prueba estadística falla, la conclusión es que la sociedad o el agente particular no se comportan como la teoría supone. La teoría en sí no es puesta en cuestionamiento.

En quinto lugar, la política económica es el «producto final» del conocimiento científico en Economía. La política económica es la ingeniería de la ciencia económica (así como la agronomía es una ingeniería de la biología y la medicina, una ingeniería de la biología y la química). Si no hubiera política económica, la Economía sería una ciencia estéril. ¿Hay alguna forma lógica para obtener proposiciones de política económica? Estas proposiciones pueden denominarse «proposiciones gama» (γ). Una teoría válida —esto es, una teoría que ha cumplido con el proceso de falsación y ha sido aceptada— debería ser la base para la política económica. Las teorías que no hayan sobrevivido el proceso de refutación, o que nunca hayan estado bajo este, no serán de ayuda.

Pero tener proposiciones beta empíricamente comprobadas no es suficiente para establecer proposiciones de política económica. No hay una línea lógica desde la teoría económica

hasta la política económica. La razón de esto es que la causalidad no implica una relación única entre variables exógenas (los medios de la política económica) y variables endógenas (los fines). Solamente si la teoría tuviera dos variables, una endógena y una exógena, podría la política económica desprenderse directamente de la teoría. Este hecho es poco común. Las teorías económicas establecen relaciones de causalidad entre muchas variables —endógenas y exógenas—. Adicionalmente, no todas las variables exógenas pueden ser usadas como medidas o instrumentos de política económica. En este caso general, el agente que define la política enfrenta varios objetivos y varios instrumentos; así, el diseño de la política económica se torna mucho más complejo.

Son necesarios entonces juicios de valor para seleccionar medios y fines. Pero estos juicios de valor no pueden derivarse lógicamente de una teoría. Este hecho es consistente con la famosa sentencia de Hume: «no se puede deducir el deber ser a partir del ser». En la ciencia, proposiciones deontológicas no pueden ser lógicamente derivadas de proposiciones ontológicas. Las proposiciones gama no se derivan lógicamente de las proposiciones alfa-beta, pero tampoco son independientes de estas.

En suma, la Economía cumple con los requisitos para ser una ciencia, y para ser una ciencia social. La compleja actividad económica puede ser reducida conceptualmente a un proceso abstracto: el proceso económico. Las variables endógenas fundamentales son la producción de bienes y su distribución entre los miembros de la sociedad. Las variables exógenas dependen de la teoría económica que se utilice. De cada teoría se puede derivar proposiciones beta, es decir, proposiciones empíricamente refutables.

La conclusión de que la Economía es una ciencia puede ser sorprendente para muchos. Paul Samuelson ha popularizado a través de su famoso texto introductorio la idea de que hasta de un loro se puede hacer economista, solo hay que enseñarle a

decir «oferta y demanda». Tres críticas serias que intentan cuestionar el carácter científico de la Economía son: (a) la Economía no puede utilizar la experimentación controlada para refutar sus teorías; (b) la Economía no puede predecir los hechos del futuro; (c) no se puede ser objetivo en las apreciaciones sobre la situación económica de un país.

La primera crítica se despeja simplemente reconociendo que el criterio de validez de una teoría es estadístico —relaciones entre los valores promedios de las variables endógenas y exógenas, en lugar de relaciones determinísticas—, y eso no requiere experimentos controlados. Si el criterio fuese la capacidad de controlar el experimento, la Astronomía no se hubiera desarrollado tanto.

La segunda crítica se elimina reconociendo la distinción que debe hacerse entre predicción y pronóstico. En efecto, la ciencia económica no puede pronosticar, no puede determinar los valores que tomarán las variables endógenas en el futuro. Pero esto ocurre en todas las ciencias. La ciencia solo puede predecir: dado los valores de las variables exógenas puede determinar el valor de las variables endógenas. Pero la ciencia no tiene capacidad para saber cuáles serán los valores de las variables exógenas en el tiempo; si quisiera conocerlo, necesitaría endogenizar las variables exógenas y construir una nueva teoría, que tampoco podría predecir, a menos que se conocieran las nuevas variables exógenas de la nueva teoría, y así entraríamos en una regresión continua. La tercera crítica no tiene que ver con la ciencia económica sino con el comportamiento de los economistas.

¿Cuál es la metodología que siguen los economistas en sus investigaciones? A juzgar por la mayoría de las publicaciones, los economistas utilizan principalmente dos métodos. Unos utilizan el método deductivo. A partir de unos supuestos determinan las relaciones que deben existir entre las variables que analizan. El uso intensivo de las matemáticas ha significado que esas derivaciones lógicas sean realizadas con mucho rigor. De

estas relaciones así deducidas se pasa, usualmente, al campo de las aplicaciones. Se supone, implícitamente, que si la derivación es lógicamente correcta, la relación propuesta debe ser también empíricamente cierta.

Otro grupo de trabajos que aparece en la literatura es de tipo más bien empírico. Dada una base de datos, los economistas derivan relaciones estadísticas entre las variables observadas. El uso intensivo de la econometría ha significado que la derivación de esas relaciones sean también realizadas con mucho rigor y sofisticación. De estas relaciones estadísticas así obtenidas se pasa, usualmente, a la causalidad. Se supone, también implícitamente, que los datos pueden contar su verdad por sí solos. Así se identifica —posiblemente seducidos por la sofisticación de los métodos econométricos utilizados— correlación o asociación con causalidad. Y sobre la base de esos resultados se pasa, comúnmente, al campo de las aplicaciones. En este caso, los economistas utilizan el método inductivo.

Al parecer, los economistas no buscan refutar las teorías que utilizan. En su famoso libro *The Methodology of Economics or How Economists Explain,* Blaug argumenta que, en la práctica, los economistas rara vez buscan refutar sus teorías. Ciertamente, los economistas están atareados masivamente con trabajo empírico, pero una gran parte de este trabajo usa la metodología del verificacionismo, concluye Blaug (1992).

Habría que añadir que una parte importante de la investigación empírica busca directamente la aplicación de una teoría, lo que se ha denominado aplicacionismo. Este método se puede encontrar en los libros de texto de Economía. Una vez que se ha presentado una teoría, sigue inmediatamente una sección en donde dicha teoría es dirigida a «aplicaciones del mundo real, de modo que pueda verse a la teoría en acción». Esta metodología supone que la teoría es consistente con la realidad, en lugar de someterla primero al proceso de falsación. Si en la aplicación la teoría no se ajusta correctamente con los datos, usualmente se introducen cambios en algunos factores no ob-

servables (preferencias, expectativas, probabilidades) para reconciliar a la teoría con los hechos. La teoría no es desafiada con el aplicacionismo; por el contrario, se la defiende.

Los economistas también usan la retórica. Esto significa que ellos se apartan de la metodología de la ciencia. En vez de ello usan el arte de la argumentación. Cualquier elemento es considerado válido para defender la teoría, incluida la teoría misma, pues esta es tomada como una verdad. Cuando los economistas defienden una teoría, ellos sostienen, implícitamente, que sus teorías no pueden ser erróneas; luego, si la realidad no coincide con la teoría, ¡algo debe estar mal con la realidad!

El aplicacionismo conduce a la concepción de que la Economía es ingeniería. En un artículo preparado para una conferencia cuyo tema era examinar si la Economía estaba llegando a ser una ciencia dura, el conocido autor de textos de microeconomía Hal Varian (1989) sostiene que es un error comparar a la Economía con la Física, mejor sería compararla con la ingeniería; de igual manera, es un error comparar a la Economía con la Biología, mejor sería compararla con la medicina. Y como no puede existir una epistemología para la ingeniería, la Economía no necesita de la filosofía de la ciencia. La Economía es aplicación de la teoría económica a problemas de política económica. La Economía es pura ingeniería.

El argumento de Varian tiene un problema lógico. No puede existir ingeniería sin una ciencia que la sustente. La Economía no puede ser ingeniería pura. En realidad, su argumento es que la ciencia económica es deductiva, y una vez creada la teoría se la debe utilizar sin necesidad de pasar por el proceso de su falsación. Su visión es que la Economía no es una ciencia fáctica sino una ciencia formal. Así, para él, la ingeniería económica se sustenta en la teoría económica (la teoría neoclásica). Esta metodología deductivista es de aceptación general entre los economistas.

La Economía se enseña a escala planetaria a través de los textos producidos en los países del primer mundo. En la indus-

tria de la educación universitaria, los textos de Economía se llaman usualmente «manuales», y son presentados como tales. El paradigma teórico que sustenta la ingeniería es la teoría que se presenta en el texto universitario. Como mencioné en la introducción, una teoría se convierte en paradigma cuando ingresa a los textos universitarios. La enseñanza de la Economía en el tercer mundo deja de preocuparse por la ciencia para convertirse en pura ingeniería.

El objetivo de la Economía es, según estos textos, la política económica. La ingeniería económica es la política económica y viceversa. De los paradigmas incorporados en los textos se extraen implicancias de política económica para cualquier realidad económica. Esta solución parece ser eficiente. Para los países pobres es más barato utilizar el conocimiento ya creado en el primer mundo y aplicarlo a sus problemas particulares en lugar de desarrollar su propia ciencia económica. Implícitamente, se acepta que la ontología universalista de la Física también se aplica a la Economía. Pero, todo esto es contrario a los fundamentos de la Economía presentados aquí.

¿Por qué buscan los economistas defender una teoría antes que refutarla? Este problema pertenece al campo de la Psicología social o la Sociología, no a la epistemología. Podría pertenecer también al campo de la Economía, en tanto exista una racionalidad económica detrás de este comportamiento. Por ejemplo, si sus salarios dependieran de las teorías que practican, los economistas escogerían y defenderían aquella que les otorga la retribución máxima.

1.4. El estudio científico del capitalismo contemporáneo

Este libro busca explicar la producción y distribución en el capitalismo contemporáneo. Las variables endógenas ya están definidas y las exógenas dependerán de la teoría que se utilice.

La metodología que se utilizará a lo largo del libro es la de alfa-beta, que incluye tres reglas básicas. En primer lugar, las teorías deben generar proposiciones empíricamente refutables. En segundo lugar, una teoría es rechazada cuando falla en todos sus modelos. Para que ello ocurra es suficiente que las proposiciones beta que se refieran a solo una variable endógena falle en todos los modelos. Y, en tercer lugar, la prueba estadística no-paramétrica tiene prioridad sobre las pruebas paramétricas.

Varias teorías serán estudiadas en este libro. Así, el próximo capítulo se dedica al estudio de las teorías que conforman la economía estándar: neoclásica, clásica y keynesiana. Debido a las dificultades que tienen estas teorías para explicar los aspectos más notables del capitalismo mundial, en los capítulos restantes del libro se presentan nuevas teorías que intentan explicar los distintos tipos de capitalismo, los del primer mundo y los del tercer mundo. Debido a que el conocimiento científico es acumulativo, el estudio de la Economía estándar se hace necesario para construir nuevas teorías. La economía estándar jugará un papel heurístico en esta tarea.

Capítulo 2
Las economías convencionales

Las teorías neoclásica, clásica y keynesiana coexisten en la ciencia económica. Dentro de este conjunto, la teoría neoclásica es la dominante y sería mejor llamarla el paradigma actual de la ciencia económica, pues esta es la teoría que se espera que todo economista sepa, sea que haya estudiado en Nueva York, Lima o Colombo. La teoría keynesiana tiene también mucha presencia en el debate entre economistas. La teoría clásica de Marx y Ricardo ha tenido una influencia muy importante en épocas pasadas, pero mucho menos hoy en día. La pregunta que se busca responder en este capítulo es la siguiente: ¿tienen estas teorías, más allá de sus grados de popularidad, capacidad para explicar las características básicas del capitalismo mundial contemporáneo?

2.1. La economía neoclásica

En esta sección, se construirá un sistema económico abstracto, la llamada economía neoclásica. Esta construcción está inspirada en la escuela iniciada por Smith (1776) y Walras (1954 —la primera publicación fue en 1883—).

Respecto al contexto institucional, en esta economía las personas participan en el proceso económico dotadas de activos económicos. Las organizaciones incluyen familias, firmas y al Gobierno. En cuanto a la racionalidad, las personas actúan guiadas por el propio interés. Ellos intercambian sus activos o los bienes producidos con dichos activos en un *mercado*. ¿Qué es un mercado? Es un mecanismo particular para la interacción social.

Las reglas para la participación de los individuos en un intercambio del mercado son las siguientes:

1. Es voluntario, dadas las restricciones en los activos económicos de los individuos y dadas las restricciones legales.
2. Es impersonal. El intercambio opera a través del precio del mercado, donde la ley de un solo precio se mantiene en el intercambio.
3. La motivación que subyace al intercambio es el interés personal de los individuos, quienes buscan tomar ventaja de cada posibilidad de intercambio para obtener ganancias en beneficio propio.

Debido a que ninguna de estas reglas es observable, debe quedar en claro que el carácter de mercado o de no-mercado de un intercambio se establece por medio de supuestos.

En la economía neoclásica, se supone que todos los mercados son *walrasianos.* Un mercado es walrasiano cuando los individuos pueden comprar o vender toda la cantidad de bienes que desean intercambiar al precio vigente de mercado. Un mercado walrasiano requiere que el precio nominal del bien —la tasa de intercambio entre el dinero y el bien— sea completamente flexible.

Los elementos esenciales de este sistema económico se presentan en el siguiente conjunto de proposiciones alfa:

$\alpha_n(1)$. *Contexto institucional.* Reglas: las personas participan en un proceso económico dotadas de activos económicos que están sujetos a derechos de propiedad privada; las personas intercambian bienes sujetos a las normas del mercado de intercambio. Organizaciones: las personas se encuentran organizadas en unidades familiares y en firmas que producen bienes; el Gobierno provee dinero y bienes públicos.

$\alpha_n(2)$. *La racionalidad económica de los agentes.* Las motivaciones de los individuos son las del *homo economicus,*

un individuo abstracto cuyas acciones están guiadas por la motivación del interés propio.

$\alpha_n(3)$. *Las relaciones del mercado.* Todos los mercados son walrasianos.

Un modelo neoclásico

Se considera una sociedad con tres mercados walrasianos: trabajo (H), un solo bien (B) y dinero (M). Se supone, además, competencia perfecta en todos los mercados. Todos los individuos son socialmente homogéneos. Todos están dotados de las mismas calificaciones o del mismo *stock* de capital humano. Sin embargo, existe un grupo que también está dotado de *stocks* de capital físico. Existe, por lo tanto, dos grupos sociales: los capitalistas y los trabajadores.

Los trabajadores intercambian sus servicios laborales por el bien B producido por los capitalistas, y los capitalistas intercambian el bien B por servicios laborales. El intercambio en el mercado se lleva a cabo mediante el uso del dinero. Los salarios son pagados al final del periodo de producción; por consiguiente, los trabajadores necesitan saldos reales para poder realizar sus transacciones. El periodo de análisis será un año; por lo tanto, la dimensión temporal de todas las variables de flujo será un año (esta especificación sobre las variables de flujo se hará en todos los capítulos de este libro). El sistema económico es estático, cerrado y libre de incertidumbre.

Cada trabajador individual está dotado de S_{hi} unidades de trabajo y S_{mi} unidades de dinero. La función de utilidad de los trabajadores tiene como argumento solamente al bien B. La cantidad de trabajo ofertada es igual, en consecuencia, a la dotación de trabajo. Los trabajadores buscan obtener la máxima cantidad del bien B, sujetos a su restricción presupuestaria y a su restricción de saldos reales.[1] Por lo tanto, podemos repre-

[1] La dimensión total del ingreso laboral se medirá en dólares/año. La tasa nominal del salario puede tener cualquier dimensión en el tiempo (ej.: dólares al

sentar el comportamiento de los trabajadores mediante el siguiente sistema de ecuaciones:

$$\begin{aligned} P_h S_{hi} &= P_b D_{bi} + (D_{mi} - S_{mi}) \\ D_{mi} / P_b &= L_i(D_{bi}),\ L_i' > 0,\ L_i'' < 0 \\ (S_{mi} - D_{mi}) / P_b &= d_{bi} \end{aligned} \qquad (2.1)$$

La primera ecuación hace referencia a la restricción presupuestaria. Esta indica que el trabajador puede gastar su ingreso en comprar el bien B y en la acumulación de saldos monetarios. En esta ecuación, P_h y P_b son los precios nominales del trabajo y del bien B, respectivamente; y D_{mi} es la cantidad de dinero demandada. Esta es una ecuación de flujo.

La segunda ecuación muestra la cantidad de dinero que requiere un trabajador para poder realizar sus transacciones, dada la tecnología de pagos de la sociedad. Esta ecuación señala que cuanto mayor es el ingreso real del individuo, mayor será la cantidad de saldos reales que debe mantener con fines transaccionales. Además, si se duplicara su ingreso real, no implicaría que se duplique su saldo real; este se incrementaría en menos del doble. Existen economías de escala en el uso del dinero. La demanda por dinero es una demanda derivada, pues es un requerimiento para hacer transacciones en el mercado, de acuerdo con las prácticas de pagos en uso en la sociedad. Esta es una relación flujo-*stock*.

La tercera ecuación muestra el mecanismo de ajuste en los saldos reales, siempre que el saldo real mantenido por el trabajador sea diferente de la cantidad demandada o requerida. Si el saldo real mantenido es mayor al requerido, el exceso se gastará en la compra del bien B; si el saldo real es menor, el trabajador debe acumular el *stock* de dinero comprando una cantidad

día/persona) en tanto que la dimensión para los servicios laborales corresponda a la misma dimensión de los servicios laborales en la oferta de trabajo (días/persona al año). La variable saldo real es un *stock*.

menor de bienes. El mecanismo de ajuste para alcanzar el saldo real requerido opera mediante el cambio en la cantidad del bien B comprado. Se supone que este ajuste en el *stock* es hecho en un periodo y llamamos d_{bi} a la correspondiente cantidad de bienes B que es intercambiada para este propósito. Por consiguiente, esta cantidad puede ser añadida o sustraída de la cantidad D_{bi} (variable de flujo) solo en el periodo de ajuste. Una vez realizado el ajuste, el trabajador comprará D_{bi} periodo tras periodo.

Dado su ingreso real, el trabajador comprará la cantidad D_{bi}. Esta cantidad determinará la cantidad de dinero D_{mi} que requiere. Dada su dotación de dinero, aparece el superávit o déficit de los saldos reales. Esta desigualdad llevará a comprar una cantidad extra del bien B para reducir los saldos reales o a comprar una menor cantidad para acumular el *stock* de dinero. Por lo tanto, en el periodo de ajuste, la cantidad comprada será mayor o menor a D_{bi} en la cantidad d_{bi}. Una vez que se ha realizado el ajuste, el trabajador comprará D_{bi} periodo tras periodo. Esta cantidad depende solo de su ingreso real.

El mecanismo para alcanzar el estado de equilibrio está dado por la tercera ecuación del sistema (2.1). La condición de equilibrio implica que $D_{mi}= S_{mi}$. Ello significa un estado donde los trabajadores mantienen sus saldos reales de manera voluntaria. Cualquier exceso o déficit de los saldos reales será ajustado intercambiando dinero por el bien B. El costo de oportunidad del dinero está determinado por el consumo del bien B.

Debido a que participa en mercados walrasianos, el trabajador puede vender toda la cantidad de trabajo que desee y comprar toda la cantidad del bien B que desee con el salario que ha ganado. Las dos primeras restricciones presupuestarias en (2.1) no son ecuaciones inocentes, ya que presuponen que los trabajadores operan en mercados walrasianos.

Un trabajador individual enfrenta dos tipos de variables: aquellas sobre las cuales no puede decidir (exógenas) y aquellas sobre las que sí puede decidir (endógenas). Dados los valores de

las variables exógenas, él escogerá el valor de las variables endógenas de acuerdo con su racionalidad económica. Estas variables son:

Variables exógenas : P_h, P_b, S_{hi}, S_{mi}
Variables endógenas : D_{bi}, D_{mi}

Dada sus restricciones, el trabajador elige la cantidad del bien B que consumirá por unidad de tiempo y el *stock* de dinero. Una vez elegidas, estas cantidades se repetirán periodo tras periodo si las variables exógenas permanecen fijas. El equilibrio estático es estable. Cualquier otro valor de D_{mi} implicará exceso o déficit en los saldos reales. Por lo tanto, se puede utilizar el método de la estática comparativa para derivar las proposiciones beta, esto es, el efecto de las variables exógenas sobre las variables endógenas.

Un incremento de P_b tendría los siguientes efectos. El ingreso real disminuiría, al igual que D_{bi}. A un nivel bajo de ingreso real, desciende el nivel requerido de los saldos reales. Por otro lado, el incremento de P_b podría llevar al trabajador precisamente a una caída del saldo real inicial. Así, ambas cantidades tenderían a igualarse. Por supuesto, si la elasticidad ingreso de los saldos reales fuese igual a uno, ambas cantidades se igualarían y el trabajador no tendría que modificar su *stock* de dinero inicial. Pero, si la elasticidad ingreso fuese menor a uno, como se supone aquí, el trabajador terminaría con un déficit en sus saldos reales. El trabajador tendría que acumular *stock* de dinero; para ello tendría que comprar una cantidad menor del bien B en el periodo de ajuste, menor a lo que indica el nuevo equilibrio. Este es el mecanismo mostrado en la tercera ecuación del sistema (2.1). Así, un cambio en el precio nominal del bien B tiene un *efecto de saldos reales* sobre la demanda del bien B. Una vez realizado el ajuste, el nuevo equilibrio implica que el trabajador podría consumir una nueva cantidad de bien B, con el nuevo saldo real, periodo tras periodo.

Si suponemos que el nivel del precio se duplica, el ingreso real del trabajador caería a la mitad; de manera similar, disminuiría su salario real. La cantidad comprada del bien B caería en la mitad, pero la cantidad de saldos reales requerida caería en menos de la mitad. El trabajador debe acumular saldos reales, y eso implica comprar menos del bien B; es decir, en el periodo de ajuste, él compra menos de la mitad del bien B. Este es el efecto saldo real que se ha originado por un cambio en el nivel del precio. Pero este efecto sería relativamente pequeño y podría darse solo en el periodo de ajuste, sin efectos en el consumo en los siguientes periodos.[2]

Consideremos, ahora, el incremento del salario nominal. La cantidad demandada del bien B aumentaría, pero el trabajador experimentaría déficit de saldos reales. El ajuste del *stock* implica comprar, en el periodo de ajuste, una cantidad del bien B mayor al del valor del nuevo equilibrio. Una vez realizado el ajuste, el nuevo equilibrio implica que el trabajador consumiría una nueva cantidad del bien B, con el uso del nuevo saldo real, periodo tras periodo.

Si el salario nominal se duplicara, la cantidad comprada del bien B también se duplicaría. La cantidad del saldo real que mantiene el trabajador no cambiaría; sin embargo, la cantidad requerida se incrementaría, aunque en menos del doble. Por lo tanto, el trabajador tendría que acumular saldos reales comprando una menor cantidad del bien B; es decir, en el periodo

[2] La ratio balance real / ingreso real es una relación flujo-*stock*. Por lo tanto, su magnitud depende de la unidad de tiempo escogida. Supongamos que el ingreso es equivalente a 100 kg del bien B por mes y el saldo real es 20 kg. Si P_b se incrementa en 20%, D_b y el saldo real caerán en 20% cada uno, hasta 80 y 16, respectivamente. Si la elasticidad ingreso es 0.8, la demanda por saldos reales caerá en 16% hasta 16.8. El trabajador tendrá una escasez equivalente a 0.8 kg del bien B. Por consiguiente, en el periodo de ajuste la demanda total será 79.2 y en los siguientes periodos será 80. Entonces, el efecto sobre el saldo real será del orden de 1% y solamente por un periodo. El menor periodo de análisis y el menor incremento ocasionan el menor efecto sobre el saldo real.

de ajuste, él compraría menos del doble del bien B. Este también es el efecto saldo real que esta vez se ha originado por un cambio en el salario nominal, que también sería pequeño y efectivo solamente en un periodo.[3]

Un incremento en la dotación de dinero generaría un exceso de saldos reales. El trabajador tendría que ajustar su *stock* de dinero y comprar una cantidad mayor del bien B en el periodo de ajuste. Pero, una vez realizado el ajuste, el equilibrio inicial prevalecerá periodo tras periodo.[4]

En suma, la cantidad demandada del bien B que proviene de los trabajadores depende solo de su ingreso real. Esta función puede ser escrita como:

$$D_{bi} = f_i (P_b, P_h, S_{hi}), \quad f_{i1}<0, f_{i2}>0, f_{i3}>0 \qquad (2.2)$$

Esta función es homogénea de grado cero en los precios nominales. Por lo tanto, puede ser escrita como:

$$D_{bi} = F_i (P_h / P_b, S_{hi}) = F_i (w, S_{hi}), \quad F_{i1}>0, F_{i2}>0 \qquad (2.3)$$

En consecuencia, dada la dotación de mano de obra del trabajador, la demanda del bien B dependerá solo de la tasa del salario real (w).

La función de la demanda agregada del bien B y la del dinero son obtenidas al sumar todas las funciones individuales de

---

[3] Usando las mismas condiciones de la nota anterior, suponemos que el salario nominal se incrementa en 20%, esto es, hasta 120 kg. La demanda por saldos reales se incrementará en 16%, esto es, en 3.2 kg. Esta escasez de saldo real implicaría una caída en del orden de 3% y solamente en el periodo de ajuste. La demanda total del bien B será de 116.8 en el periodo de ajuste y 120 en los subsiguientes periodos.

[4] Una vez más, usando las condiciones iniciales de la nota previa, suponemos que la cantidad de dinero se incrementa en 20%, esto es, 4 kg. Este exceso en el saldo monetario será usado para demandar una mayor cantidad del bien B. Esta posibilidad implica una demanda mayor en 4% y solamente durante el periodo de ajuste.

demanda de los trabajadores. No hay interacciones directas entre los trabajadores, como acciones colectivas u otras estrategias de interacción. Todos son tomadores de precio. Las interacciones entre los trabajadores solo son indirectas y se realizan a través del mercado (este supuesto sobre la agregación del comportamiento individual se aplicará en todo el libro). Por consiguiente:

$$D_b = F(w, S_h),\ F_1>0,\ F_2>0 \qquad (2.4)$$
$$D_m = P_b L(D_b),\ L'>0,\ L''<0$$

Finalmente, d_b es obtenido de la restricción presupuestaria agregada de todos los trabajadores:

$$P_b d_b = S_m - D_m = P_h S_h - P_b D_b \qquad (2.5)$$
$$P_h S_h = P_b D_b + (D_m - S_m)$$

Por otra parte, el capitalista j está dotado con K_{bj} unidades del bien B en la forma de capital físico. Su racionalidad puede ser expresada como la búsqueda de la maximización de los beneficios nominales totales (P'), que es igual a:

$$P_j' = P_b Q_{bj} - P_h D_{hj}, \qquad (2.6)$$

sujeta a la función de producción:

$$Q_{bj} = F_j(D_{hj}, K_{bj}), \qquad F_{j1}>0,\ F_{j11}<0 \qquad (2.7)$$

Para un nivel dado de conocimiento tecnológico en la sociedad, se supone que el flujo del producto por unidad de tiempo depende de la cantidad de trabajo, de los insumos de capital y de dos factores no tecnológicos: la duración de la jornada de trabajo y de la intensidad del esfuerzo que despliegan los trabajadores. Se supone que estos dos factores son dados exógenamente. Por lo tanto, el flujo del producto depende solamente de la cantidad de los insumos, como se mostró anteriormente. El producto y los beneficios son netos de depreciación.

La ecuación de la ganancia no es inocente, pues ya incorpora implícitamente el supuesto de que los capitalistas operan en mercados walrasianos. A los precios vigentes en el mercado, ellos pueden vender toda la cantidad del bien B y pueden comprar toda la cantidad de trabajo que desean intercambiar. Los capitalistas no mantienen saldos reales. Supondremos que ellos no tienen necesidad de capital de trabajo, pues pagan los salarios al final del periodo de producción. Sus beneficios no están relacionados con las tenencias de dinero; más bien, provienen del proceso de producción. Ellos obtienen sus ganancias en unidades del bien B al retener parte de su producción. Este hecho es consecuencia de la propiedad que tienen de las firmas y, por tanto, de la producción obtenida en el proceso productivo. La oferta del bien B al mercado, S_{bj}, es igual a la cantidad producida menos la cantidad retenida como beneficios.

Se supone que las firmas tienen diferentes funciones de producción. Todos conocen la tecnología actual; dicha tecnología es de rendimientos constantes a escala y de productividades marginales de los factores decrecientes. Pero hay diferencias en la productividad de las firmas: con las mismas cantidades de insumos, las firmas obtienen diferentes cantidades producidas debido a diferencias en el acceso a los recursos naturales o a la ubicación espacial. La competencia en el mercado no puede, por lo tanto, destruir los beneficios. El equilibrio del mercado puede darse con ganancias positivas.

Dado los valores de las variables exógenas, el capitalista elegirá los valores de las variables endógenas de acuerdo con su racionalidad económica. Estas variables son:

Variables exógenas : P_h, P_b, K_{bj}
Variables endógenas : $Q_{bj}, S_{bj}, D_{hj}, P$ (ganancias reales)

La condición de equilibrio de la firma es:

$$F_{j1}(D_{hj}, K_{bj})\, P_b = P_h \qquad (2.8)$$

Esta condición señala que el valor de la productividad marginal del trabajo debe ser igual a la tasa del salario nominal del mercado. La cantidad demandada de trabajo es determinada por esta condición. Esta cantidad es la que maximiza el beneficio.

Este equilibrio es estable. Si D_{hj} estuviera por encima de la cantidad de equilibrio, el valor de la productividad marginal del trabajo sería menor al salario nominal, y eso llevaría a la firma a reducir el empleo e incrementar el valor de la productividad marginal (debido al supuesto de productividad marginal decreciente). Así, se restauraría el equilibrio.

La estabilidad del equilibrio nos permite usar el método de estática comparativa para derivar las proposiciones beta:

$$D_{hj} = h_j (P_h, P_b, K_{bj}), \quad h_{j1}<0, h_{j2}>0, h_{j3}>0 \tag{2.9}$$

Esta función es homogénea de grado cero en los precios nominales. Por lo tanto, puede escribirse como:

$$D_{hj} = h_j (P_h / P_b, K_{bj}) = H_j (w, K_{bj}), \quad H_{j1}<0, H_{j2}>0 \tag{2.10}$$

La cantidad ofrecida del bien B es obtenida de la restricción presupuestaria de la firma:

$$P_b S_{bj} = P_b Q_{bj} - P'_j = P_h D_{hj} \tag{2.11}$$

Esta ecuación señala que, en equilibrio, el valor de las ventas de la firma debe ser igual al salario que se debe pagar.

Las funciones agregadas son obtenidas mediante la suma simple de las funciones individuales. La curva de la demanda agregada del trabajo es solo la suma horizontal de las curvas individuales. Dichas curvas difieren entre ellas debido a sus diferencias en la productividad. Entonces:

$$D_h = H (w, K_b), H_1<0, H_2>0 \tag{2.12}$$

La restricción presupuestaria agregada sería:

$$P_b S_b = P_b Q_b - P' = P_b Q_b - P_b Q_b + P_h D_h = P_h D_h \qquad (2.13)$$

En equilibrio, los ingresos que las firmas obtienen por sus ventas deben ser iguales al salario que deben pagar a los trabajadores. Los capitalistas retienen parte de su producto como beneficios; por lo tanto, esa parte del producto está fuera del mercado.[5]

Equilibrio general

Las condiciones del equilibrio general se pueden escribir como:

$$\begin{array}{lll} \text{Mercado laboral} & : & S_h = D_h \\ \text{Mercado del bien B} & : & S_b = D_b \\ \text{Mercado de dinero} & : & S_m = D_m \end{array} \qquad (2.14)$$

En el mercado de bienes, la demanda viene solo de los trabajadores y, por lo tanto, es igual a la masa salarial (W), esto es, D_b=W, donde W = w D_h. En consecuencia, D_b puede ser sustituido por W en la función de demanda del dinero, la segunda ecuación en el sistema (2.4).

La restricción presupuestaria agregada, que también debe ser satisfecha, es obtenida de las correspondientes restricciones de los capitalistas y de los trabajadores, es decir, de las ecuaciones (2.5) y (2.13) que luego se suman. Por eso:

$$P_h S_h + P_b S_b = P_b D_b + (D_m - S_m) + P_h D_h \qquad (2.15)$$

[5] En un mundo con muchos bienes, este supuesto es equivalente a consolidar todas las transacciones que se realicen entre capitalistas. Por consiguiente, el «mercado» se refiere aquí solamente a las transacciones entre capitalistas y trabajadores.

$$P_b(D_b - S_b) + P_h(D_h - S_h) + (D_m - S_m) = 0$$

Los trabajadores y los capitalistas interactúan a través del mercado. En esta interacción, ellos enfrentan dos tipos de variables: las que no pueden ser modificadas (exógenas) y las que sí se pueden modificar (endógenas). Dados los valores de las variables exógenas, los trabajadores y los capitalistas determinarán los valores de las variables endógenas, siendo consistentes con la restricción y la racionalidad de cada individuo. Estas variables son:

Variables exógenas : S_m, K_b, S_h
Variables endógenas : w, D_h, Q_b, P/W, P_b, P_h

Hay dos subsistemas en el sistema del mercado, representados por la ecuación (2.14): uno real (mercado de trabajo y de bienes) y uno monetario. No hay valores monetarios en el subsistema real, pero sí hay variables reales en el subsistema monetario. La solución no es simultánea sino secuencial. El subsistema monetario puede ser resuelto solamente después de que el subsistema real ha sido resuelto.

De la restricción presupuestaria agregada, ecuación (2.15), se deriva la ley de Walras. Esta ley dice que si en un sistema de mercados formados por *n* mercados walrasianos *n-1* mercados se encuentran en equilibrio, el *n*-ésimo mercado debe estar necesariamente en equilibrio. Esto implica que un mercado es redundante en la determinación del equilibrio general. Debido a que los mercados de trabajo y del bien B deben ser resueltos primero, hay un grado de libertad para elegir un mercado. Para obtener el equilibrio general es suficiente resolver primero el mercado de trabajo y luego se resuelve el mercado de dinero.[6]

El mercado laboral tiene una cantidad fija de oferta laboral; por consiguiente, el salario real será determinado por la curva

[6] Barro (1997) desarrolla un modelo similar, en el que los mercados inclui-

de demanda. Una representación gráfica de esta solución puede observarse en la Figura 2.1 (a). En este gráfico, se puede ver que el mercado del bien B también se encuentra en equilibrio, porque el área que muestra la masa salarial representa la cantidad demandada, pero también representa la cantidad ofrecida pues es igual al producto total menos los beneficios (se debe recordar que el área bajo la curva de la productividad marginal es igual al producto total). Una vez que se ha determinado la masa salarial, la demanda de saldos reales también queda determinada; además, dada la oferta de dinero, esta cantidad demandada de dinero determinará el nivel del precio. La solución del mercado de dinero se muestra en la Figura 2.1 (b).[7] De esta manera, la solución del equilibrio general es alcanzable. Los tres mercados están en equilibrio.

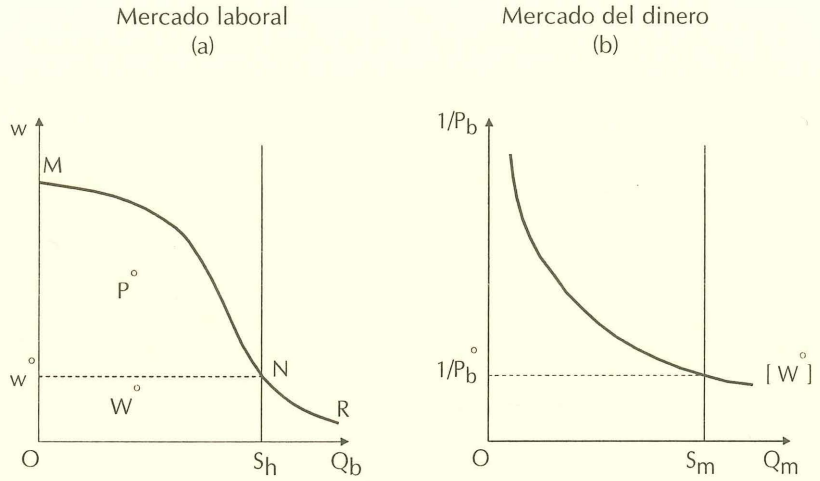
¿Cómo se llega a esta solución en esta economía abstracta? Ciertamente, no es mediante las transacciones, ni mediante mecanismos de prueba y error. Si este mecanismo fuese aplicado, las curvas de demanda podrían desplazarse después de cada ronda de transacciones, y la solución nunca sería alcanzada. En vez de ello, la teoría de Walras supone la existencia de un subastador. La solución es encontrada mediante la acción de este subastador, que cotiza precios y permite que cada individuo diga las cantidades que estaría deseoso de comprar y vender en cada mercado a esos precios. El subastador calcula el exceso de demanda y el exceso de oferta resultante en cada

---

dos son bienes, crédito y dinero. En su explicación sobre la lógica de la solución del equilibrio general, él se equivoca cuando dice lo siguiente: «nosotros podríamos obtener los mismos resultados considerando cualquiera de los pares de mercados (de los tres mercados)» (1997: 172). El punto en el que se equivoca es que, en su modelo de equilibrio general, la solución es secuencial. De hecho, él escoge el par conformado por los mercados de bienes y dinero, y muestra una solución secuencial.

[7] En la Figura 2.1 (b), tenemos en el eje vertical $1/P_b$, que es el precio del dinero en términos del bien B. Su dimensión es kg/dólares. Se debe notar también que la curva de la demanda del dinero es una hipérbola rectangular.

Figura 2.1
Equilibrio general neoclásico



mercado, y cotiza los nuevos precios, incrementando el precio donde hay exceso de demanda y reduciéndolo donde hay exceso de oferta. Estas iteraciones conducen a una solución. Una vez que la solución ha sido alcanzada, y se conocen los precios de equilibrio, todos pueden proceder a intercambiar las cantidades que establecieron. Los excesos de demanda en todos los mercados se limpian.[8]

Ciertamente no observamos que los mercados del mundo real operen con subastadores. Pero la teoría de Walras dice que cualquier mecanismo que utilice el mercado del mundo real es equivalente al del subastador, es decir, es *como si* existiera un subastador.

[8] Los subastadores no son recompensados por su trabajo. El mecanismo de limpieza del mercado es hecho a un costo mínimo.

En suma, dadas las restricciones que enfrentan los individuos, nadie tiene el poder y el deseo para cambiar esta solución. Esta es, entonces, una situación de equilibrio.

La naturaleza del equilibrio de la economía neoclásica puede ser resumida de la siguiente manera:

(a) Todos los mercados están en equilibrio, y la solución de precios y cantidades, incluida la distribución del ingreso, se repetirán periodo a periodo.
(b) Los beneficios son ingresos residuales que recibe el capitalista por su derecho de propiedad sobre el *stock* de capital. Los beneficios no constituyen el precio de bien o servicio alguno en el mercado.

En relación con el punto (b), cabe señalar que los individuos podrían haber acumulado el *stock* de capital, sea ahorrando o prestándose, posibilidad que implicaría la existencia del mercado de crédito. Aquí, el mercado de crédito ha sido ignorado. Si lo incluyéramos, habría un nuevo precio: la tasa de interés. Los individuos también podrían tomar en alquiler bienes de capital, pero entonces se tendría que incluir un mercado de alquiler de máquinas, con un nuevo precio que sería la renta. Si introdujéramos estos dos nuevos mercados en el sistema, los beneficios serían lo que le queda a la firma después del pago hecho a los trabajadores y a los servicios de crédito y de alquileres. Los beneficios no pueden ser confundidos ni con los intereses ni con las rentas. Los beneficios son ingresos residuales.

En esta economía, el producto neto que el sistema económico produce se denomina Ingreso Nacional (Y). Su distribución entre los salarios y los beneficios puede ser representada como:

$$\begin{aligned} Y &= W + P \\ &= w D_h + \kappa K \end{aligned} \qquad (2.16)$$

En la segunda ecuación, la masa salarial W es descompuesta en una tasa de salario real w y en empleo total D_h; de manera similar, los beneficios totales son descompuestos en una tasa promedio de retorno κ y en *stock* de capital K. Debe notarse que κ no es un precio de mercado alguno.

Proposiciones beta

El equilibrio general es estable y único. Estas propiedades se pueden ver en la Figura 2.1. Por lo tanto, se puede usar el método de estática comparativa para derivar las proposiciones beta. Hay tres variables exógenas en el sistema: la oferta del trabajo, el *stock* del capital y la oferta de dinero.

Un incremento en la oferta de trabajo reducirá el salario real e incrementará el empleo. Tanto el producto total como los beneficios reales se incrementarán. Sin embargo, el cambio en W es ambiguo; pues depende de la elasticidad de la curva de demanda del trabajo. El cambio en la distribución del ingreso también es indeterminado. El cambio en el nivel del precio es también indeterminado debido a que no existe un cambio definido en la curva de demanda de dinero.

Un incremento del *stock* del capital implica inversión neta en el bien B. Esta inversión, que proviene del ahorro de los capitalistas o de los trabajadores, desplazará hacia fuera la curva de demanda del trabajo. La tasa real de salarios pagados subirá, y el producto total se incrementará al igual que W. El cambio en los beneficios totales es indeterminado, al igual que la distribución del ingreso. La demanda de dinero aumenta, y eso implica que el nivel del precio cae.

Suponga ahora que el Gobierno incrementa su oferta de dinero.[9] Este incremento generará un exceso en los saldos reales

[9] El supuesto consiste en que el Gobierno inyecta dinero en la economía para mantener sin cambio las dotaciones relativas de dinero entre las personas. Este mecanismo es a veces llamado «dinero helicóptero». Bajo este supuesto, la

que mantienen voluntariamente los trabajadores. Ellos buscarán deshacerse del exceso de dinero comprando más cantidades del bien B, y eso tendería a incrementar su precio. Como resultado, el valor de la productividad marginal del trabajo tendería a incrementarse, las firmas buscarían producir más y emplear más trabajadores. Dado que todos los trabajadores están plenamente empleados, el salario nominal tendería a subir, y continuaría subiendo hasta que el salario real de equilibrio sea restaurado. El producto total se mantendría constante. Por lo tanto, el único efecto de un incremento en la oferta de dinero es el aumento de las variables nominales. No hay efecto sobre las variables reales. Este modelo predice que el dinero es neutral.

En el caso de un trabajador *individual,* si se incrementara su cantidad de dinero, él podría comprar más cantidades del bien B. Pero cuando el dinero se incrementa para *todos,* nadie comprará más cantidades del bien. Este es un caso claro del principio lógico denominado *falacia de composición:* lo que es cierto para un individuo no es necesariamente cierto para el agregado.[10] Debido a esta falacia se tendría que eliminar el efecto del saldo real en la demanda agregada para el bien B. Este resultado es lógicamente consistente con el supuesto que hiciéramos anteriormente: ignorar para el caso del individuo el efecto del saldo real, debido a que este es relativamente pequeño.

Este modelo neoclásico no predice cambios definitivos sobre la distribución del ingreso debido a los cambios en las variables exógenas. Cambios en la oferta de dinero no afectan la

política monetaria es equivalente a la política fiscal. Un incremento en la oferta de dinero implicará un incremento en el gasto del Gobierno y viceversa. Este supuesto será seguido a lo largo de todo el libro.

[10] Esta es la razón para usar oraciones en «condicional» cuando describimos el comportamiento individual (en el sentido de que es lo que le gustaría hacer al individuo), y para usar oraciones «sin condicional» cuando describimos el comportamiento agregado, esto es, la solución del equilibrio general. Es evidente que el nuevo equilibrio también es alcanzado con la ayuda de un subastador.

ratio beneficios-salarios (P/W) porque el dinero es neutral. Un incremento en el *stock* del capital tiene un efecto ambiguo sobre la distribución, porque tanto P como W aumentan, pero en proporciones relativas que son indeterminadas. Finalmente, un incremento en la oferta laboral tiene también un efecto ambiguo sobre la distribución: se incrementa P pero el cambio en W es indeterminado. Este último efecto depende de la elasticidad de la curva de demanda del trabajo, que es desconocida.

La matriz de proposiciones beta se presenta en el Cuadro 2.1. La proposición beta correspondiente a cada variable endógena aparece en la fila respectiva. La falla empírica de una fila es suficiente para rechazar el modelo. Una fila falla cuando al menos uno de los signos es contrario a la realidad. Por eso, la falla empírica de un signo en la matriz es suficiente para rechazar todo el modelo.

Cuadro 2.1
Matriz de proposiciones beta de un modelo neoclásico estático

Variables endógenas	Variables exógenas		
	S_m	K_b	S_h
w	0	+	−
D_h	0	+	+
U	0	0	0
Q_b	0	+	+
P	0	?	+
W	0	+	?
P/W	0	?	?
P_b	+	−	?
P_h	+	?	?

Consistencia empírica: los países del primer mundo

¿Se parece el mundo real a la economía neoclásica? Considérese el caso de los países del primer mundo. Ciertamente, este modelo neoclásico no constituye una descripción completa de estos países, pero —si la aproximación es lo suficientemente buena— se podría decir que estos países funcionan *como si* fuesen una economía neoclásica. Si la aproximación no es lo suficientemente buena, estos países no se parecen a la economía neoclásica y, en este sentido, se puede decir que la teoría (la abstracción) es empíricamente falsa.

Desafortunadamente, los datos empíricos sobre las variables endógenas y exógenas presentadas en la Cuadro 2.1 no están disponibles para todos los países del primer mundo. Los textos de macroeconomía contienen información pero incompleta y, usualmente, solo de la economía norteamericana. Por lo tanto, el modelo neoclásico no puede ser evaluado empíricamente; sin embargo, es posible analizar empíricamente si la economía norteamericana funciona como si fuese una economía neoclásica.

El efecto de la inversión y de población debe ser observado en un contexto de largo plazo. Las predicciones del primer efecto son consistentes con los datos empíricos. En el periodo 1960-1980, la inversión real creció en la economía norteamericana, y el crecimiento del producto estuvo alrededor del 3.6% por año, hecho que coincide con los continuos incrementos en la tasa de los salarios reales de cerca del 2.0% por año (Barro 1997, capítulo 1). La tasa de crecimiento de la población tiende a declinar en el tiempo y ayuda, probablemente, al crecimiento económico y al aumento en el salario real.

El efecto de la oferta de dinero debe ser observado en un contexto de corto plazo. Basado en datos trimestrales para el periodo 1959-1996, la economía norteamericana muestra que los cambios en la oferta de dinero son pro cíclicos (se mueven en la misma dirección que el producto), aunque moderadamente (Barro1997, capítulo 18). Esta relación estadística es inconsis-

tente con la predicción de la neutralidad del dinero del modelo. Otra inconsistencia es la existencia y la persistencia del desempleo, que es pro cíclico (Barro 1997, capítulo 6).

La no-neutralidad del dinero y la persistencia del desempleo no son características de la economía norteamericana solamente. En particular, la persistencia del desempleo es una característica evidente en el resto de países del primer mundo. También se observa de manera evidente que, en periodos de inflación, los niveles del producto total y de los salarios reales se modifican. Se puede concluir, entonces, que el modelo neoclásico presentado aquí falla en explicar el funcionamiento de los países del primer mundo.

¿Pueden otros modelos predecir diferentes resultados? Cualquier modelo de la familia neoclásica predecirá el pleno empleo del trabajo. Esto es así porque la teoría neoclásica supone que el mercado de trabajo es un mercado walrasiano. Cualquier modelo predecirá la neutralidad del dinero, pues la teoría neoclásica supone que todos los mercados son walrasianos, y eso implica que los precios nominales son flexibles (para limpiar los mercados). Así, las decisiones individuales sobre las variables reales del sistema dependen solo de los precios relativos.

En resumen, los países capitalistas no parecen asemejarse a la economía neoclásica. Ellos muestran tasas persistentes de desempleo positivo y muestran que las variables nominales y reales están conectadas. Estos datos son inconsistentes con las predicciones de todo modelo neoclásico.

Con la introducción del supuesto de información incompleta sobre los precios, se ha desarrollado una nueva teoría, llamada *teoría de las expectativas racionales*. Esta teoría señala que las personas forman sus expectativas utilizando toda la información corriente disponible. Las personas hacen uso eficiente de su información limitada para evitar cometer errores. A pesar de que las personas forman sus expectativas racionalmente, la información incompleta sobre los precios significa que un incremento no anticipado del dinero y del nivel del precio hace a los productores pensar que el precio relativo de su producto se

incrementará. Por eso, incrementos en la oferta de dinero pueden inducir a expansiones en el producto y en el empleo. Sin embargo, cambios anticipados en la oferta de dinero y en el nivel del precio no llevan a confusiones sobre el precio relativo. Por lo tanto, esta teoría puede predecir la no-neutralidad del dinero, aunque solamente para la parte no-anticipada de la oferta de dinero. Pero la distinción entre cambios anticipados y no-anticipados en la oferta de dinero no es observable, y hace que la teoría no sea falsable.

Cualesquiera que sean los méritos de los supuestos de las expectativas racionales, la teoría neoclásica emplea el supuesto de los mercados walrasianos. Este hecho lleva a que el problema del desempleo continúe sin poder explicarse. El argumento usual de que lo observado en el mercado de trabajo es solo el desempleo friccional no nos ayuda para explicar la existencia del desempleo. Aunque el subastador requiera de tiempo para limpiar el mercado, el mercado finalmente debería limpiarse. Si no es así, es porque en el mundo real el mercado laboral no funciona como un mercado walrasiano, como supone la economía neoclásica. Algunos factores esenciales en el funcionamiento del mercado laboral han sido ignorados por la teoría neoclásica.

2.2. La economía clásica

Otra economía abstracta será desarrollada en este capítulo. Esta es la llamada economía clásica. Esta construcción se inspirará en la escuela iniciada por Ricardo (1821) y Marx (1867). La estructura de la propiedad es tal que existen clases sociales en esta sociedad. La estructura del mercado incluye mercados walrasianos y no-walrasianos.

En esta sociedad hay dos clases sociales: los capitalistas dotados de capital y los trabajadores —que no tienen capital—. Los capitalistas no dan en alquiler los servicios de su capital

para evitar la formación de nuevas firmas y nuevos capitalistas. Los trabajadores pueden crear firmas solo mediante la acumulación del capital.

El sistema del mercado opera con mercados walrasianos y no-walrasianos. Un mercado es walrasiano cuando el precio de equilibrio limpia el mercado; esto es, compradores y vendedores son capaces de realizar el intercambio de los bienes en las cantidades que desean al precio del mercado. Por lo tanto, el racionamiento del mercado opera a través de los movimientos en los precios nominales. Por contraste, el mercado es no-walrasiano cuando el precio de equilibrio no limpia el mercado. En consecuencia, en este mercado, las personas no pueden llevar a cabo el intercambio de los bienes en las cantidades que están deseosos de intercambiar a los precios vigentes del mercado. El mercado opera con racionamiento de cantidades. El mecanismo del precio falla en limpiar los mercados debido a que el precio nominal o el precio relativo es determinado exógenamente.

Existe superpoblación en el mercado de trabajo. La productividad marginal de toda la fuerza laboral es menor al ingreso de subsistencia de los trabajadores. Este ingreso de subsistencia está dado exógenamente. Los capitalistas establecen el salario real del mercado al nivel del ingreso de subsistencia de los trabajadores. La determinación de este salario real implica dos cosas. En primer lugar, los trabajadores no pueden acumular capital sobre la base de su ahorro; el salario real solamente permite reproducirlos como trabajadores, con las mismas habilidades laborales en el sentido intergeneracional. En segundo lugar, a este salario real hay exceso de oferta laboral. El mercado de trabajo es no-walrasiano. Debido a que el precio relativo del trabajo (la tasa del salario real) es exógenamente determinado, el mecanismo del precio no puede limpiar el mercado de trabajo.

A causa de que los capitalistas controlan el capital y debido a la existencia de exceso en la oferta de trabajo, los capitalistas pueden establecer la jornada laboral por encima de la que hace posible producir el salario de subsistencia. La jornada laboral

deviene en un parámetro de la función de producción. Así, los servicios del trabajo utilizado por una firma tienen dos componentes: *el trabajo necesario* (la parte de la jornada laboral que sería suficiente para producir el salario de subsistencia de los trabajadores) y *el trabajo excedentario* (la parte de la jornada laboral que produce el excedente). Los capitalistas se apropian del producto producido con el trabajo excedente en forma de beneficios. Esta es la llamada *explotación*. Los beneficios provienen de la explotación. Si no hay trabajo excedentario, si no hay explotación, no hay beneficios.

Supóngase que una firma produce 100 kg de producto por trabajador en una semana —considerando que la jornada laboral dura 40 horas—; supóngase también que el ingreso de subsistencia por trabajador es igual a 50 kg. El proceso de producción puede ser conceptualmente separado en «trabajo necesario», el tiempo necesario para producir la subsistencia del trabajador, esto es, 20 horas; y el «trabajo excedente» de 20 horas que produce los 50 kg restantes para el capitalista. El trabajo necesario se transforma en salarios, mientras que el trabajo excedentario genera beneficios. Pero, entonces, la productividad del trabajo debe ser lo suficientemente alta para generar un producto por trabajador más allá del salario de subsistencia.[11] Si la productividad laboral fuese de 50 kg., esta economía no podría generar beneficios. No podría operar como una economía capitalista. Si la productividad laboral fuese menor a 50 kg., esta

[11] Roemer (1982) tiene desarrollada una teoría general de la explotación en la que el excedente laboral puede ser generado mediante el mercado de trabajo o el mercado de crédito. En este último caso, la apropiación por los capitalistas del excedente laboral opera a través del mercado de crédito y toma la forma de intereses, no de beneficios. Esta es una finura teórica; empíricamente, sin embargo, los beneficios son el componente más importante de la propiedad del ingreso en el primer mundo. Él también argumenta que la explotación por el mercado laboral necesita sobrepoblación y un sector de subsistencia. El sector de subsistencia no está introducido en este modelo porque el objetivo es explicar los países del primer mundo.

economía no podría ser socialmente viable. No podría operar bajo ningún sistema.

Los rasgos esenciales de esta economía son expresados en el siguiente conjunto de proposiciones alfa:

$\alpha_c(1)$. *Contexto institucional.* Reglas: las personas participan del proceso económico dotadas de activos económicos; dichos activos están sujetos a derechos de propiedad privada. Las personas intercambian bienes sujetos a las normas del mercado. Esto último incluye la norma de que los salarios reales son ajustados al nivel de subsistencia de los trabajadores. Organizaciones: las firmas, las familias y el Gobierno.

$\alpha_c(2)$. *Condiciones iniciales.* Los individuos están dotados de cantidades desiguales de activos económicos. Hay dos clases sociales: los capitalitas —quienes concentran el total de capital de la economía— y los trabajadores —que no poseen capital—.[12]

$\alpha_c(3)$. *La racionalidad económica de los agentes.* Las motivaciones de los agentes son las del *homo oeconomicus* Los capitalistas buscan dos objetivos particulares: la maximización de sus beneficios y la preservación de su posición en la estructura de clases. Los capitalistas no ofrecen en alquiler los servicios de sus *stocks* de capital; para ellos es preferible obtener beneficios a obtener rentas.

$\alpha_c(4)$. *Las relaciones del mercado.* Los mercados son walrasianos y no-walrasianos. Los salarios reales son ajustados al nivel del ingreso de subsistencia de los trabajadores; a este salario, se da un exceso de oferta en el merado laboral. No hay un mercado para los servicios de los bienes de capital. Los capitalistas establecen la jornada de trabajo por encima del trabajo necesario.

---

[12] Mientras que el supuesto de dos clases en la teoría neoclásica fue introducido como un supuesto auxiliar para producir un modelo particular, en la economía clásica esta estructura de clases es una proposición alfa de la teoría.

Un modelo clásico

En esta economía clásica, habrá solo tres mercados: trabajo (H), mercancía (B) y dinero (M). Los trabajadores y los capitalistas intercambian trabajo por bienes usando dinero. El mercado de trabajo es un mercado no-walrasiano; los mercados restantes son walrasianos. Existe competencia perfecta en todos los mercados. Solo los trabajadores están dotados de *stocks* de dinero que mantienen como saldos reales. El sistema económico es estático, cerrado y libre de incertidumbre. Todos los individuos (capitalistas y trabajadores) son homogéneos en términos de calificaciones laborales o capital humano. Por lo tanto, los beneficios provienen completamente de la explotación, es decir, solamente de los derechos de propiedad del capital.

Los trabajadores buscan maximizar su ingreso real, sujetos a sus dotaciones individuales de oferta laboral. Ellos intercambian su fuerza de trabajo en un mercado en el cual ellos saben que no todos los trabajadores consiguen la cantidad de trabajo que desean ofrecer. En el mercado de trabajo, existe un mecanismo de racionamiento por cantidades exógenamente determinado. Este mecanismo de racionamiento puede ser aleatorio, posibilidad que implica que las firmas, una vez que conocen la cantidad total de trabajo que quieren comprar D_h, comprarán D_{hi} unidades de trabajo de cada trabajador *i* —donde $D_{hi}= 0$ para algunos trabajadores, a la tasa del salario real dada w—.

La restricción presupuestaria para el trabajador *i* es la siguiente:

$$\begin{aligned} & w\, D_{hi} = D_{bi} + (D_{mi} / P_b - S_{mi} / P_b) \\ & D_{mi} / P_b = L_i(D_{bi}),\ L_i' > 0,\ L_i'' < 0 \\ & (S_{mi} - D_{mi}) / P_b = d_{bi} \end{aligned} \qquad (2.17)$$

Aquí $D_{hi} = n_i D_h$, donde n_i es la participación del trabajador *i* en la demanda de trabajo total. Cuando $n_i = 0$, el trabajador no está empleado. Para el comportamiento del trabajador individual, las variables relevantes son:

Variables exógenas : $w, D_{hi}, P_h, P_b, S_{hi}, S_{mi}$
Variables endógenas : D_{bi}, D_{mi}

La única condición de equilibrio para el trabajador individual *i* es asegurar que $D_{mi} = S_{mi}$. Ignorando el efecto de los cambios en la oferta de trabajo —que es nulo en una economía con exceso en la oferta laboral— e ignorando el efecto de los saldos reales, la demanda del bien B dependerá de dos variables exógenas: el salario real y la cantidad demandada de trabajo, dado el mecanismo de racionamiento. Por lo tanto, la función de demanda individual del bien B es:

$$D_{bi} = F_i(w, D_{hi}) \qquad (2.18)$$

Debido a que la suma de la cantidad comprada de trabajo hecha por las firmas a cada familia debe ser igual a la cantidad total de trabajo que están deseosas de comprar, al nivel dado del salario real, podemos mostrar que el mecanismo de racionamiento mencionado satisface esta condición:

$$\Sigma D_{hi} = \Sigma n_i D_h = D_h \Sigma n_i = D_h, \quad \text{donde } 0 \leq n_i < 1 \qquad (2.19)$$

Agregando sobre todos los trabajadores obtenemos el siguiente sistema:

$$\begin{aligned} &D_b = F(w, D_h), F_1 > 0, F_2 < 0 \\ &D_m / P_b = L(D_b), L' > 0, L'' < 0 \\ &w D_h = D_b + (D_m / P_b - S_m / P_b) \end{aligned} \qquad (2.20)$$

Sobre el comportamiento de los capitalistas, suponemos que la motivación que subyace es la de maximizar sus beneficios, sujetos a sus dotaciones individuales de capital y al conocimiento tecnológico que manejen. Para la firma *j* tenemos:

$$\text{Maximizar } P_j' = P_b Q_{bj} - P_h D_{hj}, \qquad (2.21)$$

$$\text{sujeto a } Q_{bj} = \tau f_j (D_{hj}, K_{bj})$$

Las relaciones tecnológicas que se representan en la función de producción se refieren al periodo de producción τ. Este periodo se considera como la unidad en la duración de la jornada laboral que puede ser replicada. La duración del trabajo $\tau \leq \tau'$ es exógenamente dada, donde τ' es el valor máximo que es físicamente viable.

La cantidad ofrecida al mercado es $S_{bj} = Q_{bj} - P_j$, donde P es el beneficio real. Dados los valores de las variables exógenas, el capitalista individual escoge los valores de las variables endógenas guiado por su racionalidad económica. Estas variables son:

Variables exógenas : w, P_h, P_b, K_{bj}
Variables endógenas : $Q_{bj}, S_{bj}, D_{hj}, P$

La condición de equilibrio es la siguiente: la productividad marginal del trabajo debe ser igual al salario real. Este equilibrio es estable. Se puede entonces hacer uso del método de la estática comparativa para derivar las proposiciones beta. La demanda individual de las firmas por trabajo y la oferta del bien B pueden ser representadas como:

$$D_{hj} = H_j (w, K_{bj}) \qquad (2.22)$$
$$S_{bj} = Q_{bj} - P = w D_{hj}$$

Los valores agregados son:

$$D_h = H (w, K_b), H_1 < 0, H_2 > 0 \qquad (2.23)$$
$$S_b = Q_b - P = w D_h$$

La primera ecuación es la función de la demanda de trabajo; la segunda es la oferta agregada del bien B —esta última es obtenida directamente de la ecuación del presupuesto agregado—.

Equilibrio general

Las condiciones de equilibrio para cada mercado y la restricción presupuestaria agregada son:

$$\begin{array}{lll} \text{Mercado laboral} & : & S_h = D_h + U \\ \text{Mercado del bien B} & : & S_b = D_b \\ \text{Mercado de dinero} & : & S_m = D_m, \end{array} \qquad (2.24)$$

sujeto a la restricción presupuestaria agregada:

$$w\,D_h + S_b = D_b + (D_m/P_b - S_m/P_b) + w\,D_h$$
$$(D_b - S_b) + (D_m/P_b - S_m/P_b) = 0$$

El tamaño del desempleo (U) es ahora un resultado del proceso económico. Dados los valores de las variables exógenas, la interacción entre trabajadores y capitalistas en el sistema del mercado determinará los valores de las variables endógenas del sistema. Estas variables son:

$$\begin{array}{lll} \text{Variables exógenas} & : & S_m,\ w,\ K_b,\ S_h,\ t \\ \text{Variables endógenas} & : & D_h, U, Q_b, P, W, P/W, P_b, P_h \end{array}$$

En el sistema del mercado, hay dos subsistemas, el real y el nominal; además, el sistema monetario necesita una variable real para llegar a una solución, pero no se cumple lo inverso. La solución es, por lo tanto, secuencial. La última ecuación del sistema (2.24) implica que la ley de Walras puede aplicarse solo a dos mercados. Entre el mercado de bienes y el de dinero, uno es redundante. La solución del equilibrio general implica resolver uno de los mercados walrasianos y el mercado de trabajo. Debido a que la solución debe ser secuencial, una manera de resolver el sistema es, primero, encontrar la solución en el mercado de trabajo, y, con estos valores de equilibrio, resolver el mercado monetario.

El periodo de la jornada laboral es separable en el tiempo que es necesario para producir el ingreso de subsistencia de los

trabajadores, llamado «trabajo necesario» y representado por τ^*, y el «excedente laboral» representado por τ^e. Por lo tanto, la función de producción agregada puede ser escrita como:

$$Q_b = (\tau^* + \tau^e)\, F\,(D_h, K_b), \quad \tau^* + \tau^e = \tau \tag{2.25}$$

La distribución del producto total entre los trabajadores y los capitalistas debe satisfacer las siguientes condiciones:

$$W = w\, D_h = \tau^*\, F\,(D_h, K_b) \tag{2.26}$$
$$P = \tau^e F\,(D_h, K_b)$$

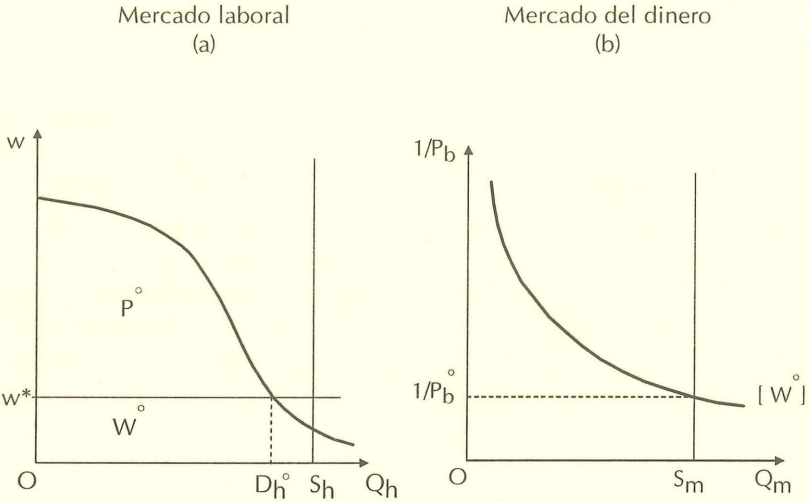
La solución del equilibrio general es presentada en la Figura 2.2. Dada la tasa del salario real, el nivel del desempleo es determinado en el mercado de trabajo. Por lo tanto, la masa salarial es conocida; dicha masa determina la cantidad demandada de dinero, y esta determina el nivel de precios. El mercado del bien está necesariamente en equilibrio, pues el área que muestra la cantidad demandada (igual a la masa salarial) también es igual a la cantidad ofrecida (el área que muestra el producto total menos los beneficios).

Una vez que precios y cantidades han sido determinados, estos valores prevalecerán periodo tras periodo. Nadie puede cambiar esta solución. Es una situación de equilibrio.

En general, la naturaleza del equilibrio en la economía clásica puede ser resumida de la siguiente manera:

(a) Los beneficios que obtiene el capitalista provienen de la apropiación del trabajo excedentario. Debido a la propiedad del capital, los capitalistas pueden emplear trabajadores y establecer la duración de la jornada de trabajo por encima del trabajo necesario. A esa apropiación se le llama explotación.
(b) La estructura social se reproducirá periodo tras periodo. Los capitalistas continuarán siendo capitalistas; los trabajadores siempre serán trabajadores. Este es el sistema social de reproducción. Los salarios reales son salarios de subsistencia, por lo

Figura 2.2
Equilibrio general clásico



que los trabajadores no van a poder acumular capital. Ellos podrían ahorrar, pero eso solamente les permitirá reproducir el capital humano que poseen en sus hijos. Los trabajadores no pueden convertirse en capitalistas con sus ahorros. Si este fuese el caso, las clases sociales desaparecerían.

El producto neto que proviene del proceso de producción se llama *excedente económico* (E). En el modelo clásico presentado aquí, los trabajadores obtienen solo salarios de subsistencia. Por lo tanto, los capitalistas se apropian de todo el excedente económico en la forma de beneficios. Así,

$$E = P = \kappa K \qquad (2.27)$$

Se debe notar que κ no es el precio de bien alguno en el mercado; es la tasa promedio de retorno del capital.

La teoría clásica puede generar modelos donde el excedente económico pueda ser apropiado en parte por los capitalistas —como beneficios— y en parte por los trabajadores —como salarios por encima del salario de subsistencia—. Los trabajadores podrían apropiarse de parte del excedente económico según su fuerza relativa en el conflicto de la distribución. Así, bajo la teoría clásica no todo el salario pagado es parte del excedente económico; este incluye solamente la parte que excede al salario de subsistencia. El salario de subsistencia es el costo de mantener el capital humano de los trabajadores, similar al costo de mantener bienes de capital en general. Este costo no es parte del producto neto.

Las proposiciones beta

En el modelo clásico presentado aquí, el equilibrio general es estable y único. Para derivar las proposiciones beta, se puede usar el método de la estática comparativa. Las variables exógenas del sistema incluyen la oferta de dinero, la oferta del trabajo, el *stock* del capital y el salario real.

Un incremento de la oferta de dinero afectará solo a las variables nominales. No habrá efecto sobre las variables reales. El modelo predice que el dinero es neutral. Un incremento en la oferta laboral afectará solo el nivel de desempleo. La inversión privada neta tendrá como efecto desplazar la curva de demanda hacia fuera, incrementando el nivel de empleo y del producto. Un incremento en el salario real exógenamente determinado tendrá como efecto la reducción del empleo y del producto, y el consiguiente incremento del desempleo.

En este modelo clásico, la distribución del ingreso debe referirse al excedente económico; sin embargo, adoptaremos el concepto de ingreso nacional, pues la separación de la masa salarial entre subsistencia y excedentaria no es posible de distinguir empíricamente. El resultado es que la distribución del ingreso nacional —la ratio beneficios-masa salarial de subsistencia— no varía

en una dirección particular como resultado de cambios en las variables exógenas. El modelo predice que cambios tanto en la oferta de trabajo como en la oferta de dinero no tendrán efectos sobre la ratio P/W. El dinero es neutral y aumentar el trabajo solamente implica mayor desempleo. Un incremento del *stock* de capital aumenta tanto P como W, pero las proporciones relativas en las cuales aumentan son indeterminadas. Finalmente, un aumento del salario real tiene un efecto ambiguo sobre la distribución, porque el cambio de W es ambiguo. La dirección de este cambio depende del valor de la elasticidad de la curva de demanda del trabajo, valor que no es conocido. La matriz de las proposiciones beta aparece en el Cuadro 2.2.

Se debería notar que, en las dos teorías presentadas —la neoclásica y la clásica—, cambios en las variables exógenas no causan cambios en dirección definida en el grado de desigualdad. En algunos casos, la desigualdad no cambia, y allí donde cambia, la dirección del cambio es ambigua. Otros modelos en cada una de las teorías tendrán las mismas predicciones. Este resultado puede ser sorprendente, pero se deriva lógicamente de las teorías.

Cuadro 2.2
Matriz de proposiciones beta de un modelo clásico estático

Variables endógenas	Variables exógenas			
	S_m	w	K_b	S_h
D_h	0	–	+	0
U	0	+	–	+
Q_b	0	–	+	0
P	0	-	+	0
W	0	?	+	0
P/W	0	?	?	0
P_b	+	?	+	0
P_h	+	?	+	0

Consistencia empírica: los países del primer mundo

Varias predicciones empíricas del modelo clásico presentado aquí no son consistentes con los hechos más resaltantes que observamos en los países del primer mundo. La predicción de la neutralidad del dinero es inconsistente con los datos empíricos. La predicción de que una inversión continua incrementará la productividad, pero mantendrá el salario real sin cambio, es también inconsistente con la experiencia de los países del primer mundo. Por ejemplo, los datos de la economía norteamericana muestran que el salario real creció a la tasa de 1.9% por año en el periodo de 1954-1973, cuando la productividad laboral ascendió a la tasa de 2.1% (Barro 1997, capítulo 6).

La teoría clásica puede dar cuenta de la existencia del desempleo, pero no puede explicar su persistencia. La inversión incrementa el empleo y el producto, pero los salarios reales permanecen fijos. El exceso de fuerza laboral, conocida en esta teoría como el «ejército industrial de reserva», impide que aparezca la escasez de trabajo y el correspondiente aumento del salario real. Pero una inversión continua podría, en última instancia, eliminar el exceso de oferta laboral.

Para evitar el agotamiento del ejército de reserva, los capitalistas podrían buscar generar nuevas tecnologías ahorradoras de trabajo. Pero esto podría requerir que los capitalistas conozcan la tecnología para generar tecnologías ahorradoras de trabajo, de forma similar al conocimiento que ellos tienen acerca de la tecnología para la producción de zapatos. No hay proposiciones alfa sobre esto en la teoría clásica. Así, la tecnología solo puede ser tomada como exógena. No hay nada en este sistema que pueda prevenir el agotamiento del ejército de reserva. Por lo tanto, esta teoría no puede explicar la persistencia del desempleo.

Cualquier otro modelo de la familia clásica generará el mismo tipo de predicciones. La razón es que la teoría supone que el salario real es exógenamente determinado. Así, el nivel de empleo depende solamente de las condiciones de la demanda

de trabajo. Pero esta demanda se expandirá en el proceso de acumulación de capital. El mismo supuesto lleva también a modelos donde el dinero es neutral. En resumen, la teoría clásica presenta dificultades para explicar el funcionamiento de las economías del primer mundo.

2.3. La economía keynesiana

Otra economía abstracta a analizar es la llamada economía keynesiana. La construcción de esta economía está inspirada en la escuela iniciada por Keynes (1936). En esta economía, no todos los mercados serán walrasianos.

Tal como se definió antes, en un mercado no-walrasiano, los individuos no pueden comprar o vender, a los precios vigentes del mercado, las cantidades que deseen intercambiar. El intercambio en un mercado no-walrasiano será llevado a cabo todavía bajo las reglas del mercado; pero habrá racionamiento en las cantidades intercambiadas. La razón, en el caso de la economía keynesiana, es que los precios nominales no son completamente flexibles y, por eso, los mercados no se pueden limpiar. ¿Por qué los precios serían rígidos en algunos mercados? Esos precios son exógenamente determinados.

La racionalidad económica de los individuos debe ser consistente con este contexto. Así, los individuos necesitan conocer el rango en el cual los precios son rígidos. Si el precio del mercado se encuentra por encima del precio walrasiano, los vendedores estarán sujetos al racionamiento. No todos los vendedores podrán vender las cantidades que desean intercambiar. Si el precio del mercado está por debajo del precio walrasiano, los compradores enfrentarán racionamiento. Los individuos conocen la situación en el tipo de racionamiento. Su racionalidad económica tomará este contexto en consideración.

En un mercado no-walrasiano, el concepto de equilibrio es igual al definido en el caso del mercado walrasiano. Los pre-

cios y las cantidades de equilibrio son aquellas que, bajo las restricciones dadas, nadie puede cambiarlo; o incluso si un individuo puede, no desearía hacerlo.

Los elementos esenciales en el funcionamiento de esta economía pueden ser presentados en el siguiente conjunto de proposiciones alfa:

$\alpha_k(1)$. *El contexto institucional.* Reglas: los individuos participan en el proceso económico dotados de activos económicos, que están sujetos a derechos de propiedad privada. Las personas intercambian bienes bajo las normas del mercado. Organizaciones: las firmas, las familias y el Gobierno.

$\alpha_k(2)$. *La racionalidad económica de los agentes.* Las motivaciones individuales son las del *homo economicus.*

$\alpha_k(3)$. *Las relaciones del mercado.* Los mercados son walrasianos y no-walrasianos. El mercado de trabajo es no-walrasiano. El salario nominal es exógenamente determinado. A este salario nominal dado, existe exceso de oferta laboral.

Un modelo keynesiano

Todos los individuos están dotados de las mismas cantidades de capital humano. Pero existen dos grupos sociales: trabajadores y capitalistas. Los capitalistas tienen, además, dotaciones de bienes de capital. Los trabajadores desean intercambiar su fuerza laboral por bienes, y los capitalistas desean intercambiar bienes por trabajo. El intercambio se realiza usando dinero, cuya oferta proviene del Gobierno.

La economía se compone de tres bienes y tres mercados: trabajo (H), un bien (B), y dinero (M). Estos mercados son de competencia perfecta. El mercado de trabajo es no-walrasiano. En este mercado, el salario nominal es dado exógenamente y fijado por encima del precio walrasiano. Los trabajadores esperan vender bajo un sistema de racionamiento. Los otros merca-

dos son walrasianos. El sistema económico es estático, cerrado y libre de incertidumbre.

La racionalidad del trabajador individual consiste en maximizar su ingreso real, la cantidad del bien B, sujeto a su restricción presupuestaria individual. Los trabajadores saben que el mercado de trabajo funciona con racionamiento de trabajo. No todos los trabajadores pueden intercambiar la cantidad de trabajo que desean vender. Una vez que las firmas tienen determinada la demanda total de trabajo D_h, al salario nominal dado, la cantidad de trabajo que los capitalistas compren de cada familia será exógenamente determinada. En el mercado de trabajo, hay un mecanismo de racionamiento dado —que supondremos es aleatorio—, tal que las firmas comprarán D_{hi} unidades de trabajo del trabajador , donde $D_{hi} = 0$ para algunos trabajadores.

Luego, para el trabajador individual i, su restricción presupuestaria puede escribirse como:

$$\begin{aligned} &P_h D_{hi} = P_b D_{bi} + (D_{mi} - S_{mi}) \\ &D_{mi} = P_b L_i(D_{bi}),\ L_i'>0,\ L_i''<0 \\ &(S_{mi} - D_{mi}) = P_b d_{bi} \end{aligned} \qquad (2.28)$$

Aquí $D_h = n_i D_h$, donde n_i está dada exógenamente y representa la participación del trabajador i en el trabajo total D_h, que es la cantidad total demandada por los capitalistas al salario nominal dado P_h. Cuando $n_i = 0$, el trabajador i está desempleado.

Dados los valores de las variables exógenas, un trabajador individual escogerá el valor de las variables endógenas de acuerdo con su racionalidad. Estas variables son:

Variables exógenas : D_{hi}, P_h, P_b, S_{hi}, S_{mi}
Variables endógenas : D_{bi}, D_{mi}

La condición de equilibrio de este trabajador es asegurar que la cantidad de saldos reales sea la que tiene que mantener en *stock*. Este equilibrio es estable.

El efecto de cambios en P_h y en P_h sobre D_b será el mismo que se da en la teoría neoclásica: el efecto de P_h será positivo y el efecto de P_b será negativo. El cambio en D_{hi} es equivalente a un cambio en la dotación laboral de los trabajadores y, por lo tanto, su efecto será positivo. Así, si se ignora el efecto de cambios en la oferta de trabajo —efecto nulo en un mercado con exceso de oferta laboral—, y el efecto del saldo real —desde que este efecto es pequeño—, obtenemos la siguiente función de demanda para el bien B:

$$D_{bi} = F_i(P_b, D_{hi}, P_h),\ F_{i1}<0,\ F_{i2}>0,\ F_{i3}>0 \qquad (2.29)$$

La demanda agregada para el bien B se obtiene sumando todas las funciones individuales de demanda de los trabajadores:

$$D_b = S\ D_{bi} = F(P_b, D_h, P_h),\ F_1<0,\ F_2>0,\ F_3>0 \qquad (2.30)$$

Debido a que la suma de las compras de trabajo hecha por las firmas a cada familia debe ser igual al total de la demanda laboral , podemos mostrar que el mecanismo de racionamiento que se ha supuesto aquí satisface tal condición:

$$\Sigma\ D_{hi} = \Sigma\ n_i D_h = D_h \Sigma\ n_i = D_h, \quad \text{donde } 0 \leq n_i < 1 \qquad (2.31)$$

La demanda total de trabajo D_h aparece en la función de demanda del bien B, porque el ingreso total nominal de los trabajadores depende del salario nominal y de la cantidad demandada de trabajo. En un mercado walrasiano, en el que los trabajadores pueden vender toda su oferta laboral, el ingreso total de los trabajadores depende solamente del salario nominal.

La función de demanda F es homogénea de grado cero en los precios nominales. Pero, debido a que el salario nominal es fijo, esta función no puede ser escrita en términos del salario real. No hay un significado económico en la afirmación «si

ambos precios nominales se duplican, la cantidad demanda permanecerá sin cambio», pues el salario nominal es fijo.

La demanda agregada de dinero es obtenida de manera similar.

$$D_m = P_b L(W),\ L'>0,\ L''<0$$
$$W = w D_h = (P_h / P_b) D_h \qquad (2.32)$$
$$D_m = M(P_b, D_h, P_h),\ M_1>0,\ M_2>0,\ M_3>0$$

En el agregado, la demanda total del bien B es igual a la masa salarial (W). Esta demanda puede, a su vez, ser expresada en términos de los componentes de la masa salarial. Esta es la formulación de la función de la demanda de dinero presentada arriba.

Los valores de las derivadas parciales en la tercera ecuación de (2.32) se han obtenido bajo el supuesto de que la elasticidad-ingreso de la demanda del dinero —la función L— es menor a uno. Luego, si el nivel del precio se incrementa, los trabajadores requerirán un mayor saldo nominal para transar la misma cantidad de bienes. Pero, al mismo tiempo, el salario real caerá y, así, el ingreso real de los trabajadores disminuirá; ellos requerirán ahora, un saldo real menor. El efecto neto será positivo si el efecto directo es mayor al efecto indirecto. Esto ocurrirá si la elasticidad-ingreso de la demanda del dinero es menor a uno. El efecto directo significa un incremento proporcional en los saldos reales, y el efecto indirecto —debido a este supuesto sobre la elasticidad— significa una disminución menos que proporcional. Los efectos de P_h y D_h son claramente positivos, pues ambas variable exógenas incrementan el ingreso real de los trabajadores.

La restricción presupuestaria de todos los trabajadores es:

$$P_h D_h = P_b D_b + (D_m - S_m) \qquad (2.33)$$

El ingreso nominal total de los trabajadores es igual a la tasa del salario nominal por la cantidad total del trabajo demandado, no

a la cantidad ofertada, como en el caso del mercado de trabajo walrasiano.

Las dos funciones de la demanda de trabajo y la oferta del bien B serán iguales a las que fueron derivadas en el caso de la teoría neoclásica. La única diferencia es que estas funciones no pueden ser expresadas en términos del salario real, porque la tasa del salario nominal es fija. Por lo tanto, para el agregado de las firmas surge el siguiente sistema:

$$D_h = H(P_b, P_h, K_b),\ H_1>0,\ H_2<0,\ H_3>0 \qquad (2.34)$$
$$P_b S_b = P_h D_h$$

La primera ecuación muestra la curva de demanda del trabajo expresada en términos del precio nominal. Esta no puede ser expresada en término del precio relativo (el salario real), porque el salario nominal está dado exógenamente. La proposición de que una caída en el nivel del precio podría generar una caída en el salario nominal para restaurar el salario real de equilibrio no se puede mantener como cierta. La función de la demanda del trabajo es homogénea de grado cero en los precios nominales, pero el mercado no opera de esta forma. La segunda ecuación muestra la restricción presupuestaria de las firmas. La oferta del bien B es obtenida de esta ecuación.

Equilibrio general

Las condiciones que la solución del equilibrio general debe satisfacer son las siguientes:

$$\begin{array}{lll} \text{Mercado laboral} & : & S_h = D_h + U \\ \text{Mercado del bien B} & : & S_b = D_b \\ \text{Mercado de dinero} & : & S_m = D_m, \end{array} \qquad (2.35)$$

sujeto a la restricción presupuestaria agregada:

$$P_h D_h + P_b S_b = P_b D_b + P_h D_h + (D_m - S_m)$$
$$P_b (D_b - S_b) + (D_m - S_m) = 0$$

Dados los valores de las variables exógenas, la interacción entre los capitalistas y los trabajadores en el mercado determinará los valores de las variables endógenas del sistema. Estas variables son:

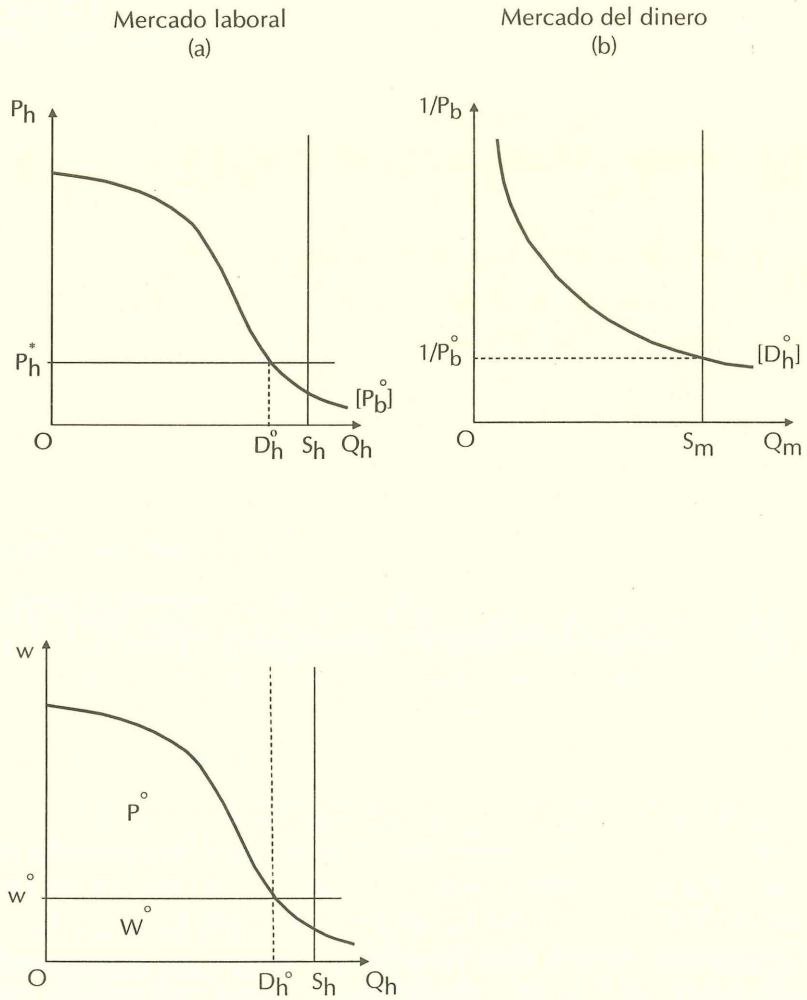
Variables exógenas : S_m, P_h, K_b, S_h
Variables endógenas : w, D_h, U, Q_b, P, W, P/W, P_b

En la búsqueda de la solución para el sistema (2.35), observamos que aquí no existen subsistemas. La solución es simultánea. La ley de Walras se aplica solo a los mercados walrasianos. En los mercados no-walrasianos, las cantidades ofrecidas y demandadas son idénticas, y son eliminadas de la restricción presupuestaria agregada. De los dos mercados walrasianos, uno es redundante. Por lo tanto, el sistema debe resolver un mercado walrasiano y el mercado no-walrasiano. Aquí se consideran los mercados de trabajo y de dinero. Estos dos mercados, representados por dos ecuaciones —(2.32) y la primera ecuación del sistema (2.34)—, pueden resolver las dos variables endógenas P_b y D_h.

Gráficamente, los mercados de trabajo y de dinero se muestran en la Figura 2.3. El sistema debe ser resuelto para P_b y D_h. Pero, para resolver D_h debemos conocer P_b, y para conocer el precio necesitamos conocer D_h. La solución es simultánea. La solución en el mercado de trabajo depende de la solución en el mercado de dinero y viceversa. El equilibrio en el mercado de trabajo implicará que el salario real esté por encima del salario walrasiano y en la curva de demanda del trabajo.

Tenemos, así, una solución estática de equilibrio general. Este proceso de producción y de distribución se repetirá periodo tras periodo. Nadie puede cambiar esta solución. Es una situación de equilibrio.

Figura 2.3
Equilibrio general keynesiano



El equilibrio keynesiano es un equilibrio con desempleo. El salario real resultante estará por encima del salario walrasiano en la curva de demanda pero no en la curva de oferta. El brazo relevante en el mercado laboral es la curva de demanda.

Algunas veces, el equilibrio keynesiano es visto a la luz de la teoría neoclásica. Los capitalistas están sobre su curva de demanda para el trabajo y sobre la curva de oferta del bien B. A los precios de equilibrio, ellos pueden realizar el intercambio de la cantidad que desean. Con los ingresos de sus ventas, ellos pagan los salarios en efectivo. Pero los trabajadores no se encuentran sobre su curva de oferta de trabajo ni sobre su curva de demanda para el bien B. A los precios de equilibrio, los trabajadores desean vender más trabajo y con esa ganancia quisieran comprar una mayor cantidad del bien B. En resumen, hay un exceso de oferta en el mercado de trabajo y un exceso de demanda en el mercado de bienes. Y esto es consistente con lo que dice la ley de Walras. «Los "precios de equilibrio" no son tales» es lo que la teoría neoclásica podría decir.

La confusión no es difícil de señalar. La curva de demanda para el bien B en el argumento neoclásico es una suerte de «demanda nocional»: ¿cuál sería la demanda para el bien B si los trabajadores vendieran toda su oferta laboral? Esto supone ya un equilibrio general neoclásico. En el equilibrio keynesiano, los trabajadores tienen una curva de «demanda efectiva» del bien B, determinada por el ingreso que ellos pueden conseguir en el mercado de trabajo. Este salario en efectivo es igual al valor de la oferta de los bienes. Así, hay un equilibrio en el mercado de bienes. El único mercado no-walrasiano, que opera con exceso de oferta, es el mercado de trabajo.

El producto neto proveniente del proceso económico se llama *ingreso nacional*. Su nivel Y y su distribución pueden ser representados de la siguiente forma:

$$\begin{aligned} Y &= w D_h + P \\ &= w D_h + \kappa K \end{aligned} \qquad (2.36)$$

Como antes, hay que notar que no es un precio del mercado sino la tasa de retorno promedio del *stock* de capital K.

Proposiciones beta

El equilibrio general es estable. Este resultado puede verse en la Figura 2.3, en la que el equilibrio se presenta en términos de los dos mercados solamente. En el mercado walrasiano, otro valor de P_b diferente al de equilibrio podría crear un exceso de demanda o de oferta de dinero y, entonces, este precio tendría que ajustarse restaurando el equilibrio. En un mercado no-walrasiano, la condición de estabilidad es trivial. Si las firmas han empleado trabajadores en una cantidad diferente a la del valor de equilibrio, ellos restaurarán el equilibrio sin ningún problema de coordinación. La estática comparativa puede ser aplicada para obtener las proposiciones beta del modelo. Debido a que los resultados de la estática comparativa no son fáciles de determinar gráficamente, debemos encontrarlos matemáticamente.

El efecto de la oferta monetaria se puede encontrar diferenciando en el sistema —dado por las ecuaciones (2.32) y (2.34)— y en las ecuaciones del mercado del trabajo y del dinero con respecto a S_m. Esto nos da el siguiente sistema:

$$\begin{aligned} \partial D_h/\partial S_m - H_1\ \partial P_b/\partial S_m &= 0 \\ M_2\ \partial D_h/\partial S_m - M_1\ \partial P_b/\partial S_m &= 1 \end{aligned} \qquad (2.37)$$

La solución de este sistema implica que $\partial D_h/\partial S_m > 0$, y también que $\partial P_b/\partial S_m > 0$. Un incremento en la oferta monetaria aumenta tanto el nivel de precio como el empleo; debido a que el salario nominal es fijo, el salario real caería. Lo que pasa con la masa salarial es ambiguo. Debido a que el empleo es mayor, el producto total se incrementará. Los beneficios también se incrementarán. El dinero no es neutral. Parte del efecto va hacia la cantidad del empleo y parte al nivel del precio.

El efecto del salario nominal puede ser encontrado diferenciando las ecuaciones del mercado de trabajo y de dinero en el sistema de equilibrio general respecto a P_h. Haciendo esto, obtenemos el siguiente sistema:

$$\begin{aligned} \partial D_h/\partial P_h - H_1\, \partial P_b/\partial P_h &= H_2 \\ M_2 \partial D_h/\partial P_h + M_1 \partial P_b/\partial P_h &= -M_3 \end{aligned} \qquad (2.38)$$

La solución de este sistema implica que $\partial D_h / \partial P_h < 0$, mientras que el signo de $\partial P_b / \partial P_h$ es ambiguo. Un aumento en el salario nominal llevará a un nuevo equilibrio, en el que el nivel del empleo es menor. Este hecho podría reducir la demanda por dinero y aumentar el nivel del precio. Pero, al mismo tiempo, un incremento del salario nominal aumentaría directamente la demanda por dinero. Por lo tanto, el efecto neto sobre el nivel del precio es indeterminado. La caída en el empleo implica un aumento en la tasa del salario real. El cambio en la masa salarial es incierto. La caída en el empleo también implica que el producto total cae. Los beneficios también caen.

El efecto de la acumulación del capital sobre el equilibrio general puede ser encontrado diferenciando ambas ecuaciones con respecto a K_b, operación que trae como resultado el siguiente sistema:

$$\begin{aligned} \partial D_h/\partial K_b - H_1\, \partial P_b/\partial K_b &= H_3 \\ M_2\, \partial D_h/\partial K_b + M_1 \partial P_b/\partial K_b &= 0 \end{aligned} \qquad (2.39)$$

La solución nos da que $\partial D_h / \partial K_b > 0$ y $\partial P_b / \partial K_b < 0$. Un incremento en el *stock* del capital expandirá la curva de demanda del trabajo. Debido a que hay desempleo, más trabajadores podrían ser empleados al mismo salario real. La masa salarial, entonces, se incrementaría. Este hecho nos llevaría a un incremento en la demanda por dinero. Dada la oferta del dinero, el nivel del precio caería. Este hecho implica un aumento en el salario real.

Así, las firmas emplearían una mayor cantidad de trabajadores, pero no tanta como la que querían en un comienzo. Por lo tanto, el nuevo equilibrio general implicará un mayor salario real y un bajo nivel de precio. Finalmente, un incremento en la oferta laboral no tendrá efecto sobre las variables endógenas, exceptuando el de un mayor desempleo.

El modelo keynesiano no predice cambios definitivos en la distribución del ingreso como resultado de cambios en las variables exógenas. Un incremento en la oferta laboral no tiene efecto sobre la ratio P/W. En esta economía, mayor oferta laboral solo implica mayor desempleo. Un incremento en el *stock* del capital aumenta tanto P como W, pero en proporciones que no son conocidas. Un aumento en cada una de las variables nominales —oferta de dinero y la tasa de salario nominal— tiene efectos ambiguos sobre la distribución del ingreso, porque el efecto sobre W es indeterminado. Las proposiciones beta de este modelo se presentan en el Cuadro 2.3.

Cuadro 2.3
Matriz de proposiciones beta de un modelo keynesiano estático

Variables endógenas	Variables exógenas			
	S_m	P_h	K_b	S_h
w	–	+	+	0
D_h	+	–	+	0
U	–	+	–	+
Q_b	+	–	+	0
P	+	–	?	0
W	?	?	+	0
P/W	?	?	?	0
P_b	+	?	–	0

Consistencia empírica: los países del primer mundo

La ausencia de la neutralidad del dinero que se observa en los países del primer mundo es consistente con el modelo de la teoría keynesiana presentado aquí. El modelo también predice una relación positiva entre inversión y producto, pero además una relación negativa entre producto-inflación en un contexto de crecimiento económico. Basado en los datos publicados por el Fondo Monetario Internacional y el Banco Mundial en sus reportes anuales, para el periodo 1968-1997, a este autor le fue posible calcular los coeficientes de correlación de Spearman entre la tasa de crecimiento del producto y la tasa de inflación para cinco países. Estos coeficientes son los siguientes: Japón -0.83, Norteamérica -0.66, Reino Unido -0.65, España -0.52 y Alemania -0.50. Todos estos coeficientes son estadísticamente significativos (en una prueba no paramétrica). Esta muestra podría ser considerada como representativa de los países del primer mundo. Estos coeficientes negativos son consistentes con la predicción del modelo keynesiano.

La existencia del desempleo también es consistente con la predicción del modelo keynesiano. Sin embargo, esta teoría tiene problemas lógicos importantes para explicar la *persistencia* del desempleo. En primer lugar, el salario nominal no es verdaderamente exógeno, pues tiene que suponerse que está fijado por encima del salario walrasiano para que pueda generar desempleo. Así, el salario nominal no es independiente del valor del salario walrasiano. En segundo lugar, un incremento continuo en la oferta de dinero o en la inversión puede hacer desaparecer el desempleo y, así, la economía keynesiana se convertiría en una economía neoclásica. En la práctica, se invoca a la teoría keynesiana cada vez que los Gobiernos desean llevar a la economía al pleno empleo. El desempleo es, según esta teoría, un problema circunstancial y soluble en la economía capitalista; es decir, el desempleo no es una característica estructural del capitalismo.

Ningún modelo de la familia keynesiana llegará a diferentes conclusiones. Los modelos con funciones de ahorro y de inversión, o con incertidumbre —tan popular en los modelos keynesianos recientes—, no producen resultados diferentes respecto de la función del mercado de trabajo. La metodología alfa-beta dice que es suficiente que la teoría falle en una proposición beta —el desempleo en este caso— para concluir que la teoría tiene dificultades en explicar todo el sistema económico. Este es el mayor problema de la teoría keynesiana para explicar el funcionamiento de los países del primer mundo.

Este capítulo ha mostrado que las tres teorías convencionales —neoclásica, clásica y keynesiana— fallan en explicar la existencia o la persistencia del desempleo. La teoría neoclásica no predice desempleo. Las teorías clásica y keynesiana predicen un equilibrio con desempleo, pero el desempleo no cumple ninguna función en el funcionamiento de la economía, pues puede existir equilibrio con pleno empleo. En el caso de la economía keynesiana, el equilibrio con pleno empleo puede ser alcanzado a través de políticas estatales.

Para estas teorías, el desempleo no es una característica estructural del capitalismo, presupuesto que es inconsistente con la observación de que los países del primer mundo *siempre* operan con desempleo. Como observa el historiador económico Garraty (1978): «El desempleo es la enfermedad del capitalismo». Se requiere de una teoría que explique la persistencia del desempleo y el papel que juega el desempleo en el funcionamiento de la economía capitalista. El capítulo siguiente da inicio a esa tarea.

Capítulo 3
La sociedad épsilon

Cualquier sistema teórico que intente explicar la producción y la distribución en el primer mundo debe contener, en primer lugar, una teoría del mercado de trabajo en la que el desempleo tenga un papel fundamental para el funcionamiento del sistema. En segundo lugar, debe mostrar que existe interacción entre variables reales y monetarias, tal como la evidencia empírica lo muestra. En este capítulo se construye una sociedad abstracta —denominada economía épsilon— en la cual se cumplirán esas dos condiciones.

3.1. Una sociedad de clases sociales

En la sociedad épsilon, existen dos clases sociales: los capitalistas y los trabajadores. En esta sociedad abstracta, los trabajadores y los capitalistas no son socios perfectos. Los capitalistas buscan extraer el esfuerzo máximo posible de los trabajadores con el fin de maximizar ganancias. Los trabajadores desean ganar el salario real más elevado efectuando el menor esfuerzo. Este conflicto social surge por la alienación de los trabajadores de la propiedad de la firma. Por lo tanto, los capitalistas necesitan implementar mecanismos para extraer el mayor esfuerzo de los trabajadores.

En esta sociedad, los trabajadores se opondrán a cualquier intento de reducir su salario nominal como una forma de evitar la reducción de su salario real. Esta resistencia está fundamentada en las normas sociales: es considerado socialmente inaceptable reducir los salarios nominales. La reducción de los salarios

reales vía inflación es vista como un resultado impersonal del funcionamiento del sistema económico y puede ser tolerada. El salario nominal puede elevarse si las firmas consideran rentable hacerlo; la rigidez del salario se da hacia abajo, no hacia arriba.

Para la economía épsilon, las proposiciones básicas son las siguientes:

$\alpha_\varepsilon(1)$. *Contexto institucional.* Reglas: las personas que participan en el proceso económico están dotadas de activos económicos; dichos activos están sujetos a derechos de propiedad privada. Las personas intercambian bienes según las reglas del mercado; estas reglas incluyen la norma que señala que el salario nominal no se puede disminuir. Organizaciones: familias, firmas y el Gobierno.

$\alpha_\varepsilon(2)$. *Condiciones iniciales.* Los individuos están dotados de cantidades desiguales de activos económicos. Existen dos clases sociales: los capitalistas y los trabajadores.

$\alpha_\varepsilon(3)$. *Racionalidad económica de los agentes.* Las motivaciones individuales son las del *homo economicus.* Los capitalistas tienen dos objetivos particulares: mantener su posición social y maximizar sus beneficios. En el mercado de trabajo, los trabajadores buscan maximizar su salario y minimizar su esfuerzo. Los capitalistas emplean mecanismos para extraer esfuerzo de los trabajadores.

$\alpha_\varepsilon(4)$. *Relaciones de mercado.* Hay mercados walrasianos y no-walrasianos. El mercado de trabajo es no-walrasiano. No existe el mercado de alquileres de servicios de bienes de capital. Los capitalistas no alquilan sus bienes de capital porque ellos prefieren las ganancias a las rentas.

3.2. La naturaleza del mercado de trabajo

En los últimos años, se han desarrollado nuevas teorías acerca del mercado de trabajo para explicar la existencia y persis-

tencia del desempleo. Una de ellas es la teoría de la búsqueda. Esta teoría supone demanda y oferta, pero además supone que la información en el mercado de trabajo es costosa. Por lo tanto, habrá trabajadores en busca de empleo y firmas en busca de trabajadores. Al mismo tiempo, las firmas despedirán a los trabajadores que no cumplan con la disciplina laboral establecida. Si las tasas de despidos y las tasas de nuevos empleos son constantes, entonces la economía tenderá a una tasa de desempleo constante. Esta es la llamada «tasa de desempleo natural» (Barro 1997).

A pesar de la existencia de los costos de información, en el largo plazo, el mercado laboral se tendría que limpiar. Le tomará más tiempo al subastador limpiar el mercado, pero deberá finalmente limpiarse. No hay nada en el sistema que obstaculice este resultado. En el largo plazo, la tasa natural de desempleo sería igual a cero. La teoría de la búsqueda no explica por qué, dados los costos de información, el mercado de trabajo finalmente no se limpia. Esta teoría tiene una inconsistencia lógica para explicar la existencia y persistencia del desempleo: introduce costos de información en el mercado laboral pero mantiene a la vez el supuesto de que este mercado es walrasiano.

La teoría desarrollada por Lucas (1981) sostiene que todo desempleo observado es voluntario. Lucas argumenta que si un trabajador está desempleado (buscando un empleo) se debe a que para él es más rentable estar desempleado que tomar cualquier otro empleo de salario más bajo. Por lo tanto, el desempleo es voluntario, es producto de una elección económica. Siempre existirá la opción de trabajar en cualquier cosa en lugar de estar desempleado. Si el salario esperado por el trabajador cuando busca empleo en *su* mercado de trabajo es mayor que la alternativa de trabajar en otro mercado de trabajo, él decidirá seguir buscando empleo (y continuará desempleado). Un ingeniero preferirá seguir buscando empleo como ingeniero antes que tomar un empleo de albañil. En este sentido, el desempleo del ingeniero es voluntario, pues podría tomar el

empleo de albañil, el empleo que podríamos llamar «empleo lucasiano».

Este comportamiento supone que el mercado de los albañiles es walrasiano, mientras que el mercado de trabajo de los ingenieros es no-walrasiano. De forma más general: todos los mercados de trabajo son walrasianos excepto el mercado de trabajo particular en el que el individuo normalmente trabaja. Este argumento tiene un problema de consistencia lógica, pues cae en el problema de la falacia de composición, es decir, puede ser cierto para el trabajador individual pero no puede ser cierto para el agregado de trabajadores.

En adición, la teoría de Lucas no es falsable. Si se observa a un trabajador en busca de un empleo, se concluirá que su salario esperado es relativamente elevado y que está voluntariamente desempleado; y si se le observa en un «empleo lucasiano» (trabajando en cualquier otro empleo de menor calificación), se concluirá que el salario esperado en su principal ocupación es relativamente bajo. Haga lo que haga el trabajador, la teoría nunca falla. La introducción de la hipótesis de expectativas, en donde no se sabe que variables observables determinan esas expectativas, convierte a la teoría en inmortal.

La teoría del salario de eficiencia es otro intento de explicar la existencia y la persistencia del desempleo. El supuesto básico dice que los capitalistas usan el salario real como un mecanismo de disciplina laboral. La intensidad del trabajo es un elemento de la función de producción; y esta intensidad es endógena, pues depende del salario real.

Inicialmente, esta teoría tenía como supuesto que la intensidad del trabajo dependía solamente del salario real, donde esta relación es positiva y tiene la forma de S. Este supuesto implica que se tenga una curva en forma de U invertida entre el esfuerzo por dólar gastado en salarios y la tasa salarial. El salario real que da el nivel más alto de ganancias para cualquier cantidad de trabajo empleada es aquel que corresponde al punto más elevado de la curva. Este es el salario de eficiencia, que determina la intensidad óptima de trabajo (Romer 1996).

En este modelo, la relación entre esfuerzo y tasa salarial es la misma para todas las firmas. Por lo tanto, el salario de eficiencia y la intensidad de trabajo óptima toman los mismos valores para todas las firmas. Pero entonces, el salario real de mercado es determinado antes de que se haya dicho nada acerca de las cantidades transadas. El salario real es exógeno.

Dado que la intensidad de trabajo es parte de la función de producción (para cada cantidad de insumos el producto a obtener depende del grado de esfuerzo de los trabajadores), el nivel de la curva de la productividad marginal (la curva de demanda de trabajo) depende de la intensidad del trabajo de los trabajadores. Una vez que la intensidad de trabajo óptima ha sido determinada, la curva de demanda de trabajo de la firma también queda determinada. Dada esta curva, la cantidad demandada de trabajo será la que maximice las ganancias, donde el salario real sea igual a la productividad marginal del trabajo.

Al salario de eficiencia exógenamente determinado, cada firma escoge la cantidad demandada de trabajadores, y esto sucede para todas las firmas, y así se determina la cantidad demandada en el mercado. Pero estas cantidades demandadas pueden ser mayores, iguales o menores que la cantidad ofrecida. Solo en este último caso se tendrá un caso de desempleo. Por lo tanto, la teoría del salario de eficiencia puede conducir a un equilibrio del mercado de trabajo con desempleo, pero este resultado no es una necesidad. El desempleo se daría solamente por accidente.

Además, hay un problema de consistencia lógica con esta teoría. Si el salario de eficiencia coincide con el salario walrasiano, en un mercado de trabajo competitivo, un trabajador que es despedido podrá encontrar otro empleo con el mismo salario. El costo de incurrir en indisciplina laboral sería nulo para el trabajador. Este costo sería significativo si la tasa de desempleo en el mercado de trabajo fuese positiva. Por lo tanto, la teoría de los salarios de eficiencia requeriría de la existencia de desempleo. Este modelo de la teoría de la eficiencia no puede, en consecuencia, explicar el desempleo.

De hecho, han surgido otros modelos de la teoría de salarios de eficiencia. Se supone que la intensidad del trabajo depende positivamente de la tasa de salario real y de la tasa de desempleo (Shapiro y Stiglitz 1984, Bowls 1985). Para un nivel de desempleo dado, a mayor salario real mayor será el costo de la indisciplina laboral y, por lo tanto, mayor el esfuerzo que ponga el trabajador en la firma. Para un salario real dado, a mayor nivel de desempleo mayor será la dificultad de conseguir empleo y mayor el costo de la indisciplina laboral y, entonces, mayor el esfuerzo que ponga el trabajador en la firma.

Para una intensidad de esfuerzo deseada por la firma, existen combinaciones diferentes de tasas de salario real y niveles de desempleo que fuerzan al trabajador a cumplir con el grado de intensidad de trabajo deseado por la firma. Con el fin de garantizar esta intensidad de trabajo, la tasa salarial debe incrementarse cuando el desempleo cae. Cuando el grupo de desempleados se vuelve más pequeño, es más fácil encontrar trabajo y, consecuentemente, una tasa salarial real más elevada disuade a los trabajadores de relajar la disciplina laboral.

Existe entonces una relación positiva entre el salario real y el nivel de empleo que asegura un grado de intensidad de trabajo dado. Esta relación se llama «la curva de extracción del esfuerzo». Dado este nivel de intensidad de trabajo, la curva de demanda de trabajo para cada firma queda determinada. Agregando todas las curvas de demanda de las firmas, se obtiene la curva de demanda del mercado.

Hay ahora tres variables endógenas a ser determinadas en el mercado de trabajo: la tasa salarial, el nivel de empleo y el nivel de desempleo. Además, hay tres ecuaciones: la demanda, la oferta y la curva de extracción del esfuerzo. Esta última es de pendiente positiva y se aproxima asintóticamente a la curva de oferta. El equilibrio del mercado laboral se da por la intersección entre la curva de la demanda y la curva de extracción de esfuerzo.

Los desempleados desearían estar empleados al salario de mercado vigente y estarían dispuestos a hacer el esfuerzo re-

querido. La competencia entre los trabajadores debería reducir el salario; sin embargo, no puede haber equilibrio a un salario más bajo. Las firmas no lo aceptarían porque saben que, si contratan más trabajadores a salarios más bajos, la intensidad de trabajo decrecerá. Nadie puede cambiar esta solución. Hay un equilibrio con desempleo. El mercado de trabajo es no-walrasiano.

Este modelo de la teoría de los salarios de eficiencia presenta, sin embargo, una dificultad de diferente naturaleza. La solución en el mercado de trabajo es independiente de las variables monetarias, depende solo de variables reales. Integrando este mercado de trabajo en un sistema de equilibrio general, la teoría predice la neutralidad del dinero. Como se dijo anteriormente, cualquier teoría que predice neutralidad del dinero falla en explicar los rasgos básicos de los países del primer mundo.

Se necesita una teoría nueva del mercado de trabajo. En la sociedad épsilon, las relaciones laborales son de conflicto. Los capitalistas buscan obtener el esfuerzo máximo de sus trabajadores a la menor tasa salarial posible. Los trabajadores buscan obtener el salario más alto con el menor esfuerzo posible. Esta es la naturaleza del mercado de trabajo. El supuesto básico es que el desempleo, no la tasa de salario real, es el mecanismo que los capitalistas emplean para obtener la disciplina laboral. Un modelo simple de lo que podría denominarse la *teoría del desempleo de eficiencia* se presentará aquí. Se supone, como hasta ahora, competencia perfecta en el mercado laboral y en los mercados de bienes.

La intensidad de esfuerzo de los trabajadores (λ) es introducida explícitamente en la función de producción. Este es un factor que determina el nivel de producción, dada las cantidades de trabajadores y capital, y dada la duración de la jornada de trabajo. Para la firma *j,* la función de producción para producir el bien B puede ser descrita como:

$$Q_{bj} = \lambda F_j (D_{bj}, K_{bj}) \qquad (3.1)$$

Aquí, D_h es la cantidad de trabajo demandada en el mercado de trabajo y K_b es la dotación de capital de la firma.

La intensidad de trabajo depende del costo que el trabajador tendrá que pagar si es encontrado en falta en el lugar de trabajo. Este costo dependerá de la posibilidad de ser detectado, que, a su vez, dependerá de los recursos para la supervisión usados por la firma (asumidos aquí como un costo fijo) y la pérdida de ingreso que el trabajador sufrirá si es despedido de su empleo. Si existiese siempre pleno empleo, es decir, si el mercado de trabajo fuese walrasiano, este costo sería cero. Los trabajadores que fuesen encontrados en falta y fuesen despedidos, encontrarían siempre otro trabajo. Por lo tanto, no habrían incentivos para la disciplina en el trabajo. El mercado laboral no podría operar como un mercado walrasiano. La naturaleza del mercado de trabajo es distinta a la naturaleza del mercado de la papa, que se utiliza siempre como el paradigma del mercado walrasiano.

Para los trabajadores, el costo de la indisciplina laboral será más elevado cuanto más elevado es la tasa de desempleo (u) y cuanto más elevado es el salario real del mercado (w). Pero, dadas las condiciones de la demanda y la oferta del mercado laboral, la tasa de desempleo y la tasa salarial no son independientes una de la otra; más aún, existe una relación de uno a uno (y positiva) entre la tasa de desempleo y la tasa de salario real. Estos supuestos se pueden escribir como sigue:

$$\begin{aligned} &\lambda = f(u,w), \text{ donde } f_1>0 \text{ y } f_2>0 \\ &w = h(u), \text{ donde } h'>0 \\ &\text{Luego, } \lambda = g(u), \text{ donde } g'>0 \end{aligned} \qquad (3.2)$$

Por lo tanto, en su forma reducida, el grado de intensidad de trabajo depende positivamente de la tasa de desempleo.

Con el fin de simplificar el modelo aún más, se considerarán solo dos niveles de esfuerzo, tales como:

$$\lambda = \lambda^* \text{ si } u \geq u^* \text{ y } w \geq w^*$$
$$\lambda = \lambda' < \lambda^* \text{ si } u < u^* \text{ y } w < w^* \qquad (3.3)$$

Aquí, u^* es el umbral de tasa de desempleo por encima del cual los trabajadores cumplen cabalmente con la disciplina laboral exigida (λ^*) y por debajo del cual no ponen todo el esfuerzo (λ'). Este umbral es socialmente determinado. El grado de conflicto en las relaciones laborales se supone que es un factor determinante, y se lo considera exógeno. El desempleo es un mecanismo para obtener disciplina laboral y, de este modo, extraer las mayores ganancias de la intensidad de esfuerzo de los trabajadores. El término w^* es la tasa salarial consistente con la tasa de desempleo u^* en el mercado de trabajo. Así, w^* representa también un valor de umbral.

La firma individual presenta dos curvas de productividad marginal de trabajo, una para cada nivel de esfuerzo. Estas pueden ser representadas por las siguientes ecuaciones:

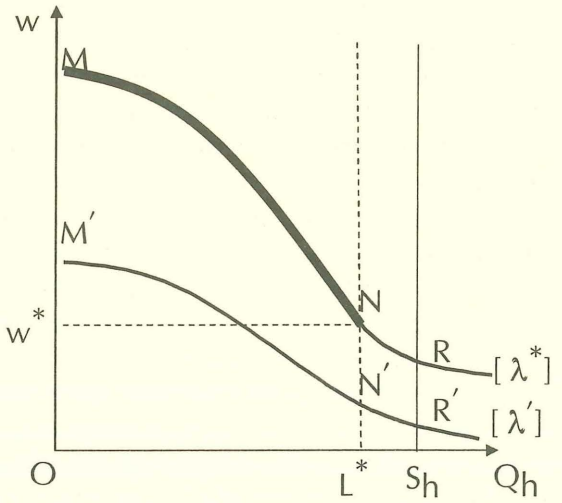
$$\lambda^* F_{j1}(D_{hj}, K_{bj}) \text{ si } u \geq u^* \text{ y } w \geq w^* \qquad (3.4)$$
$$\lambda' F_{j1}(D_{hj}, K_{bj}) \text{ si } u < u^* \text{ y } w < w^*$$

Sumando sobre todas las firmas, se determina la curva agregada de productividad marginal del trabajo. Esto se aprecia en la Figura 3.1. La curva MNR representa la productividad marginal del trabajo cuando el grado de intensidad de trabajo es igual a λ^*. La curva M'N'R' representa la productividad marginal del trabajo cuando el grado de intensidad de trabajo es igual a λ'. Por lo tanto, la curva superior requiere que la tasa de desempleo sea igual o mayor que el valor umbral u^*; si la tasa de desempleo es menor que este valor, entonces la curva relevante sería la inferior.

Supóngase que la oferta de trabajo está exógenamente dada y que toma el valor de S_h. Dado el valor de u^*, se determina el nivel de empleo máximo que puede mantener la intensidad de trabajo óptima. Sea L^* este nivel de empleo máximo, como se

aprecia en la Figura 3.1. El segmento MN es la productividad marginal de trabajo relevante cuando el empleo es igual o menor que OL^*. Si el nivel de empleo fuese más allá de L^*, la productividad marginal de la curva se trasladaría hacia la curva M'N'R'.

Figura 3.1
Eficiencia y esfuerzo laboral



¿Dónde está situada la curva de demanda de trabajo ahora? Para el nivel de salario w^*, la cantidad de trabajo demandada será igual a L^*. En niveles de salario real por encima de w^*, las firmas demandarán las cantidades indicadas por el segmento MN de la curva superior de productividad marginal (se refiere a la curva MNR). En niveles por debajo del salario real w^*, la curva de demanda no podría ser el segmento NR. ¿Por qué? Porque significaría que el empleo total estuviese más allá de L^*, y la tasa de desempleo podría tomar un valor por debajo de u^*. La intensidad de trabajo disminuiría y la curva de productividad marginal se desplazaría hacia abajo, la curva M'N', y eso implicaría un nivel de ganancia menor que el obtenido al salario w^*.

Las firmas no tendrían, en consecuencia, incentivo alguno para pagar salarios reales por debajo de w^*. Si, por alguna razón, el salario real del mercado estuviese por debajo de w^*, las firmas tendrían incentivos para pagar salarios reales mayores y desplazar la curva de productividad marginal del trabajo hacia arriba. El salario real w^* opera como un precio piso, como un umbral de eficiencia salarial. Por lo tanto, en la curva de productividad marginal de trabajo MNR, solamente el segmento MN representa la curva de demanda de empleo.[1]

De acuerdo con esta teoría, el mercado de trabajo no puede operar como un mercado walrasiano. La solución no puede ocurrir en el punto R de la Figura 3.1, donde las curvas de demanda y oferta se cruzan. Observando la figura, parece que las firmas pudiesen obtener mayores ganancias en el punto R comparado con el punto N. Pero esto es un espejismo. Al nivel del salario walrasiano, la curva de productividad marginal del trabajo se desplazaría hacia abajo situándose en M'N'R', y las ganancias serían menores que en el punto N.

Si el mercado laboral no puede eliminar los excesos de oferta (no se limpia), entonces es un mercado no-walrasiano. Por otro lado, los mecanismos de racionamiento que hemos supuesto son los de dejar en la situación de desempleo a un grupo de trabajadores elegidos de manera aleatoria (no hay nada personal en el racionamiento). Se puede, entonces, concluir que según esta teoría, *el mercado laboral funciona con exclusión.* En suma, los capitalistas usan el desempleo como una medida de disciplina para los trabajadores. Esta racionalidad implica que existe un umbral para los valores de desempleo (u^*) y salario

[1] Las ganancias totales podrían ser mayores de las que se obtienen en el punto N en la Figura 3.1 si el salario real pudiese tomar valores por debajo de w^* pero manteniendo el empleo total constante en L^*. Sin embargo, esta solución requeriría una acción colectiva de los capitalistas que, por el supuesto del modelo, está descartado. Por lo tanto, el segmento NL^* no es parte de la curva de demanda de trabajo.

real (w^*), umbrales que imponen límites a la determinación del equilibrio en el mercado laboral.

Esta conclusión implica que la teoría del desempleo de eficiencia y la teoría de los salarios de eficiencia son complementarias, una necesita de la otra. Por lo tanto, ellas pueden ser consideradas como dos modelos de una teoría más general, que puede llamarse la *teoría de la extracción del esfuerzo laboral.* Debido a que el modelo de la teoría del desempleo de eficiencia es consistente con el supuesto de que los salarios nominales son rígidos hacia abajo, este es el modelo que se empleará en el resto de este estudio.

3.3. Un modelo estático de la economía épsilon

Para desarrollar un modelo de la economía épsilon se introducirán varios supuestos auxiliares. Se supone que todos los individuos están dotados de la misma cantidad de capital humano, pero existe un grupo que está, además, dotado de capital físico. La economía produce un solo bien, el bien B. Los trabajadores desean intercambiar su trabajo por el bien B, y los capitalistas desean intercambiar el bien B por trabajo. El intercambio es realizado a través del dinero, que es provisto por el Gobierno. Existen tres mercados: el mercado de bienes (B), el mercado de trabajo (H) y el mercado de dinero (M). Estos mercados exhiben competencia perfecta. La economía es estática, cerrada y libre de incertidumbre.

La racionalidad del trabajador individual consiste en la búsqueda de la maximización del consumo del bien B, sujeto a su restricción presupuestaria y a la restricción de mantener saldos reales de dinero. Los trabajadores conocen que el mercado de trabajo funciona con exclusión. No todos los trabajadores serán capaces de intercambiar las cantidades de trabajo que están deseosos de vender. Una vez que las firmas han determinado la cantidad total demandada, la cantidad de trabajo que los capi-

talistas compran de cada trabajador es exógenamente determinada. El mecanismo de racionamiento es aleatorio.

Entonces, para cada trabajador individual *i*, su comportamiento estará representado por el siguiente sistema de ecuaciones:

$$\begin{aligned} &P_h D_{hi} = P_b D_{bi} + (D_{mi} - S_{mi}) \\ &D_{mi} = P_b L_i(D_{bi}),\ L_i' > 0.\ L_i'' < 0 \\ &S_{mi} - D_{mi} = P_b d_{bi} \end{aligned} \qquad (3.5)$$

La primera ecuación muestra la restricción presupuestaria del trabajador, que expresa que el trabajador puede gastar su ingreso nominal ($P_h D_{hi}$) en la compra del bien B (D_{bi}) y en la recomposición de sus saldos reales (D_{mi}-S_{mi}). Aquí, D_{hi} es la cantidad de trabajo que los capitalistas compran del trabajador *i*, tal que $D_{hi} = n_i D_h$ —donde n_i es la participación del trabajador *i* en el empleo total—. El trabajador *i* está desempleado cuando $n_i = 0$. La segunda ecuación se refiere a la restricción de los saldos reales: cuanto mayor es el ingreso real del trabajador, mayor será la cantidad de saldos reales de dinero que debe mantener para propósitos de hacer transacciones. La tercera ecuación se refiere al mecanismo de ajuste para restaurar la cantidad de saldos reales requeridos. Este ajuste implica $D_{mi} = S_{mi*}$

Dado su ingreso real, el trabajador comprará D_{bi} unidades de B. Esta cantidad determinará los saldos reales requeridos por las transacciones. Dada su dotación inicial de dinero, se producirá un desajuste en los saldos reales requeridos. El trabajador hará el ajuste comprando una mayor o menor cantidad de bien B en la cantidad d_{bi}. Este ajuste ocurrirá en un solo periodo. Después de que este ajuste se realice, el trabajador mostrará una situación de equilibrio en su intercambio en el mercado. La cantidad D_{bi} será comprada periodo tras periodo hasta que las variables exógenas cambien. Las variables pueden ser clasificadas como:

Variables exógenas : P_b, P_h, S_{mi}, S_{hi}, D_{hi}
Variables endógenas : D_{bi}, D_{mi}, d_{bi}

La condición de equilibrio $D_{mi} = S_{mi}$ es estable, y la estática comparativa puede ser aplicada. La cantidad demandada del bien B depende positivamente de P_h y de D_h, y negativamente de P_b. Dado que las decisiones de los trabajadores se toman de manera independiente (no existe acción colectiva entre los trabajadores), la función de demanda agregada es obtenida de la suma de las funciones individuales. Todos son tomadores de precios en el mercado. Estas funciones son las siguientes:

$$D_b = F(P_b, D_h, P_h),\ F_1 <0,\ F_2>0,\ F_3>0$$
$$D_m = P_b L(D_b),\ L'>0,\ L''<0 \qquad (3.6)$$
$$P_h D_h = P_b D_b + (D_m - S_m)$$

La cantidad de trabajo demandada D_h es incluida en la función de demanda de bienes porque el ingreso nominal total de los trabajadores depende de esta cantidad y no solo de los salarios nominales. Para los trabajadores, D_h es exógeno (en un mercado walrasiano donde los trabajadores pueden vender toda su oferta de trabajo, su ingreso total dependerá solamente del salario nominal)

Sobre la lógica del capitalista, se supone que el capitalista individual *j* está dotado con K_{bj} unidades de capital. Su racionalidad de maximización de ganancias puede presentarse como la búsqueda de maximizar las ganancias nominales P_j' de la siguiente manera:

$$\text{Maximiza}\ \ P_j' = P_h Q_{bj} - P_h D_{hj}, \qquad (3.7)$$

$$\text{sujeto a}\ \ Q_{bj} = \lambda F_j(D_{hj}, K_{bj}),\ \text{donde}\ F_{ji}>0\ \text{y}\ F_{jii}<0,\ i=1,2$$
$$\lambda = g(u),\ \text{donde}\ u \geq u^*,\ \text{tal que}\ \lambda = \lambda^*$$

Se supone que las firmas tienen diferentes funciones de producción debido no a diferencias en el conocimiento de la tecnología sino a ciertas diferencias en la localización de las firmas, tales como diferencias en el acceso a las distintas calidades de los recursos naturales y a los bienes públicos. La tecnología es de rendimiento a escala constante y está sujeta a rendimientos marginales decrecientes. Los capitalistas obtienen ganancias en unidades del bien B guardando parte de la producción. La cantidad del bien B ofertada en el mercado es S_{bj}, y es igual a la cantidad producida menos la cantidad retenida como ganancias, y la cantidad ofertada es vendida en el mercado al precio P_b.

La condición de equilibrio de la firma es la siguiente:

$$\lambda^* P_b F_{j1} (D_{hj}, K_{bj}) = P_h \qquad (3.8)$$
$$\lambda^* F_{j1} (D_{hj}, K_{bj}) = w, \text{ donde } w \geq w^*$$

La primera condición dice que el valor de la productividad marginal del trabajo debe ser igual al salario nominal. La segunda condición dice que la productividad marginal del trabajo debe ser igual al salario real, sujeto a la restricción de que este salario no sea menor que el salario de eficiencia w^*. Si el salario real está por debajo de w^*, el capitalista sabrá que los trabajadores disminuirán significativamente la intensidad de su trabajo. Entonces, no pagará salarios inferiores a w^*. La demanda de la curva de trabajo será igual a un segmento de curva solamente, y no a la curva entera de la productividad marginal del trabajo.

Para los valores dados de las variables exógenas, el capitalista individual escoge el valor de la variable endógena de acuerdo con su racionalidad. Estas variables son clasificadas como sigue:

Variables exógenas : P_h, P_b, K_{bj}
Variables endógenas : $Q_{bj}, S_{bj}, D_{hj}, P_j', P_j$

La condición de equilibrio es claramente estable. Las proposiciones beta pueden ser derivadas y son expresadas de la siguiente manera:

$$D_{hj} = H_j(P_b, P_h, K_{bj}),\ H_{j1}>0,\ H_{j2}<0,\ H_{j3}>0 \qquad (3.9)$$
$$S_{bj} = Q_{bj} - P_j'/P_b = D_{hj} P_h/P_b$$

El efecto de P_b en la demanda de trabajo es positivo porque aumenta el valor de la productividad marginal del trabajo; el efecto de P_h es negativo porque aumenta el costo del trabajo. Un incremento en el *stock* de capital tendrá un efecto positivo en la cantidad demandada de trabajo debido a que incrementará la productividad marginal del trabajo. La segunda ecuación muestra que la cantidad ofrecida del bien B es igual a la masa salarial real. Esta ecuación puede ser expresada en términos nominales como una restricción presupuestaria de la firma: el valor de sus ventas debe ser igual a la masa salarial nominal.

Las funciones agregadas son obtenidas por la sumatoria simple sobre todas las funciones individuales. Se supone que los capitalistas toman sus decisiones de manera independiente, no existe entre ellos acción colectiva alguna ni otras estrategias de interacción. Todos son tomadores de precio. Luego, en el agregado se tiene:

$$D_h = H(P_b, P_h, K_b),\ H_1>0,\ H_2<0,\ H_3>0 \qquad (3.10)$$
$$P_b S_b = P_b Q_b - P' = P_h D_h$$

La primera ecuación muestra la función de demanda de trabajo, que es homogénea en grado cero en los precios nominales. Sin embargo, no puede ser escrita en términos de precios relativos (salarios reales) debido a que P_h no es plenamente flexible. La segunda ecuación presenta la cantidad ofrecida S_b, que se obtiene de la restricción presupuestaria agregada de las firmas. El ingreso que las firmas reciben de las ventas debe ser igual a la masa salarial. Esta igualdad implica que, en términos reales,

la cantidad ofrecida del bien B debe ser igual a la masa salarial real (el salario real multiplicado por el nivel de empleo).

3.4. Equilibrio general

Existen tres mercados en la economía: el de trabajo, el de bienes y el de dinero. El mercado de trabajo es no walrasiano, y allí U representa el total de desempleo; los dos últimos mercados son walrasianos. Las condiciones de equilibrio general incluyen:

$$\begin{aligned} &\text{Mercado de trabajo} : S_h = D_h + U, \text{ sujeto a } P_h \geq P_h^* \text{ y } u \geq u^* \\ &\text{Mercado de bienes} : S_b = D_b \\ &\text{Mercado de dinero} : S_m = D_m, \end{aligned} \tag{3.11}$$

sujeto a la restricción presupuestaria agregada:

$$P_b(D_b - S_b) + (D_m - S_m) = 0$$

Para valores dados de las variables exógenas del sistema, la interacción entre los capitalistas y los trabajadores en el mercado determinará los valores de las variables endógenas del sistema. La distribución inicial en la dotación individual de activos —denominada de ahora en adelante como variable δ— debe ser incluida entre las variables exógenas. Las variables del sistema teórico serán clasificadas de la siguiente manera:

$$\begin{aligned} &\text{Variables exógenas} : S_m, K_b, S_h, \delta \\ &\text{Variables endógenas:} D_h, P_b, P_h, w, D_b, S_b, D_m, u, U, P/W \end{aligned}$$

El salario nominal P_h es endógeno, pero está restringido a tener valores por encima de su valor histórico P_h^*; es decir, no es un precio totalmente flexible. Lo mismo puede decirse de las variables desempleo (u) y salario real (w).

La ecuación de la restricción presupuestaria agregada indica que uno de los mercados walrasianos es redundante (debido a la ley de Walras). Lo que las firmas gastan debe ser igual al ingreso de los trabajadores; además, lo que los trabajadores gastan debe ser igual al valor de ventas de las firmas si los saldos reales de los trabajadores no varían. Si los saldos reales están en equilibrio, entonces, los ingresos y gastos deben ser iguales en el mercado de bienes.

Con el fin de determinar el equilibrio general en los mercados, el sistema necesita resolver solo un mercado walrasiano y un mercado no-walrasiano. Se considerarán el mercado de dinero y el mercado de trabajo. Luego, las siguientes dos ecuaciones resolverán las dos variables endógenas P_b y D_h:

$$S_m = P_b L(D_b) = M(P_b, D_h, P_h),\ M_i > 0, \text{ para todo } i$$
$$D_h = H(P_b, P_h, K_b),\ H_1 > 0,\ H_2 < 0,\ H_3 > 0 \qquad (3.12)$$

Una vez que dichas variables endógenas se determinan, se podrá conocer el resto de variables endógenas del sistema.

La interacción del mercado de trabajo y el mercado de dinero se muestra en la Figura 3.2. El panel (a) presenta el mercado de trabajo en términos de salarios nominales, donde P_h^* es el valor del salario nominal históricamente dado. El panel (c) presenta el mismo mercado de trabajo pero en términos de salario real. El panel (b) muestra el mercado de dinero. El equilibrio general resulta de la interacción entre el panel (a) y el panel (b). La curva de demanda de trabajo queda determinada una vez que el valor de P_b es conocido; asimismo, la curva de demanda de dinero queda determinada una vez que el valor de D_h es conocido. Por lo tanto, se necesita conocer el valor de P_b para determinar D_h en el mercado de trabajo, y se necesita conocer el valor de D_h para determinar P_b en el mercado de dinero; es decir, D_h depende de P_b, y P_b depende de D_h. La solución es claramente simultánea. Dado el valor del salario nominal, esta solución determina el salario real.

La solución de equilibrio general puede ser vista en la Figura 3.2. El nivel de empleo de equilibrio es D_h^o y el nivel de precio de equilibrio es P_b^o. Esta solución implica que la tasa de desempleo de equilibrio es mayor que u^*. El salario real de equilibrio es w^o, que está por encima de w^*.

La propiedad del equilibrio general de que también el mercado de bienes está en equilibrio se puede ver en el panel (c). La cantidad demandada es igual a la masa salarial. La cantidad ofrecida es igual al producto total menos las ganancias, y esta área es igual al área que mide la masa salarial, es decir, la cantidad demandada. El panel (c) muestra también la distribución de ingresos de equilibrio.

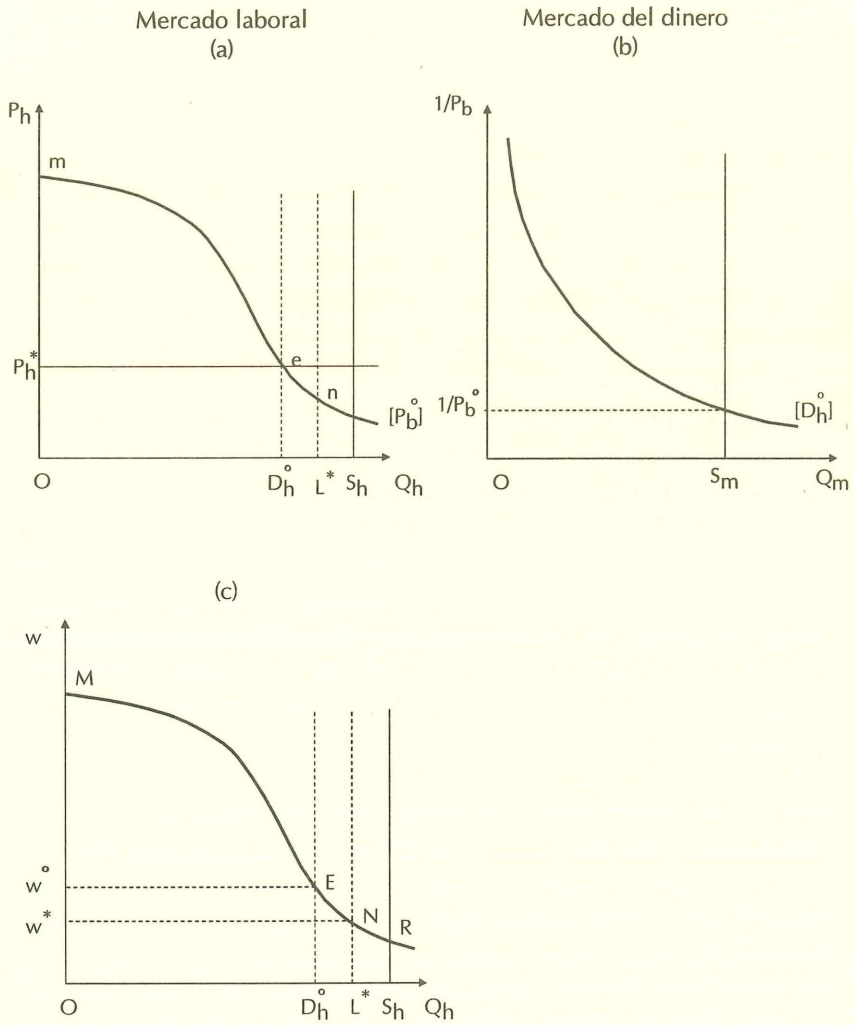
Podría parecer que las ganancias son mayores en el punto R, a la tasa de salario real walrasiano. Esto es una apariencia, debido a que este punto no es accesible. Al salario real walrasiano, la curva entera se desplazaría hacia abajo y la productividad marginal del trabajo no alcanzaría el punto R sino algún punto por debajo. Como resultado, las ganancias caerían. Entonces, la curva de demanda de trabajo es solo el segmento MN. Las ganancias son mayores en el punto N que en el punto E, donde se requiere que el salario real sea w^*, pero esta situación ocurrirá si se da una caída en el salario nominal P_h^*, y eso está descartado por los supuestos del modelo. Entonces, el equilibrio se da en el punto E, ya que ningún agente tiene el poder de cambiar esta situación.

En suma, el equilibrio general se da con desempleo. El desempleo es una necesidad para el funcionamiento de la economía épsilon. Esta economía no podría funcionar sin desempleo.

3.5. Proposiciones beta

El equilibrio general establecido es estable y único. Este hecho se aprecia en la Figura 3.2. Si el nivel de precios fuese mayor que el nivel de equilibrio, la curva de demanda de traba-

Figura 3.2
Equilibrio general con desempleo



jo se desplazaría hacia arriba y el nivel de empleo se incrementaría. Como resultado, la curva de demanda de dinero se desplazaría hacia arriba y el nivel de precios caería, restaurándose el nivel de precios de equilibrio. Entonces, la estática comparativa puede ser usada para derivar proposiciones beta.

Debido a que el equilibrio general puede darse con una tasa de desempleo mayor o igual a la tasa de desempleo de eficiencia u^*, hay dos casos que deben ser considerados en el uso de estática comparativa. Se considerará, en primer lugar, el caso en el que el equilibrio inicial se da cuando $u>u^*$.

El efecto de un cambio en la oferta de dinero en las variables endógenas puede ser hallado usando el método de la estática comparativa. Ordenando las derivadas parciales del sistema (3.12) en la forma de un sistema de ecuaciones, se tiene:

$$\begin{aligned} M_1\ \partial P_b/\partial S_m + M_2\ \partial D_h/\partial S_m &= 1 \\ H_1\ \partial P_b/\partial S_m - \quad \partial D_h/\partial S_m &= 0 \end{aligned} \qquad (3.13)$$

La solución de este sistema implica que $\partial P_b/\partial S_m > 0$ y $\partial D_h/\partial S_m > 0$. El resultado de la solución de tal sistema es que la mayor oferta de dinero tiene un efecto positivo en el valor de equilibrio del nivel de precios y del nivel de empleo. Más dinero se destina, en parte, a incrementar el nivel de precios y, en parte, a incrementar el empleo (y la producción). El dinero es no-neutral.

Las consecuencias sobre los valores de equilibrio en las otras variables endógenas son claras. El salario real cae debido al incremento en el nivel de precios. El total del producto se eleva debido al incremento en el nivel de empleo. Las ganancias se elevan debido a que el desempleo se incrementa y el salario real cae. El cambio en la masa salarial es ambiguo y, por lo tanto, el cambio en la distribución de ingresos (P/W) también lo es.

Un incremento en el *stock* de capital desplazará la curva de demanda de trabajo hacia arriba. El cambio tomará diferentes formas según el tipo de retornos a escala de la tecnología: constante, creciente o decreciente. El caso de los retornos a escala

decrecientes será ignorado, debido a que la situación puede evitarse por el mecanismo de replicación del tamaño óptimo de planta. Los retornos crecientes a escala también serán ignorados debido a que este tipo de tecnología no es lógicamente consistente con el universo de un bien que es intercambiado en un mercado de competencia perfecta. Por consiguiente, los retornos constantes a escala constituyen el supuesto que es consistente con el modelo teórico.

Bajo esta tecnología, la curva de demanda de trabajo se desplaza hacia fuera, pero no de manera paralela. La nueva curva comienza en el mismo intercepto y va separándose cada vez más de la curva inicial. En el punto donde los insumos de trabajo y de capital se duplican, el producto (el área bajo la nueva curva) se duplicará, y la curva pasará a través del intercepto del nivel de empleo que es el doble y por el nivel inicial de salario real. Esto cumple con la propiedad de que, bajo una tecnología de retornos constantes a escala, la productividad marginal del trabajo permanece constante si la ratio de los insumos permanece constante.

El sistema de ecuaciones (3.12) se somete ahora al mismo método anterior y se obtiene el siguiente sistema para resolver el efecto del aumento en el *stock* de capital:

$$\begin{aligned} M_1\, \partial P_b/\partial K + M_2\, \partial D_h/\partial K &= 0 \\ H_1\, \partial P_b/\partial K - \partial D_h/\partial K &= -H_3 \end{aligned} \qquad (3.14)$$

La solución de este sistema implica que $\partial P_b/\partial K < 0$ y $\partial D_h/\partial K > 0$. Por lo tanto, un incremento en el stock de capital tiene el efecto de incrementar el nivel de empleo y disminuir el nivel de precios. Este último efecto es el resultado de un incremento en la demanda de dinero y de una cantidad dada de oferta de dinero que hace que la gente incremente sus saldos reales y, por lo tanto, reduzca la compra del bien B.

El efecto sobre los valores de equilibrio de las variables endógenas restantes también queda determinado. Los salarios

reales se incrementan. El producto total también se eleva. En términos de distribución, la masa salarial se eleva. El cambio en las ganancias es ambiguo debido a que el salario real se incrementa. Si el salario real permaneciera constante las ganancias se elevarían. Luego, el cambio en la ratio P/W es ambiguo. El ajuste en el mercado de trabajo no requiere ningún cambio en el salario nominal. El efecto de un incremento en el *stock* de capital en el equilibrio del mercado de trabajo se puede ver gráficamente en la Figura 3.3. Un incremento en el *stock* de capital de K_0 a K_1 tiene el efecto de elevar tanto el salario real como el nivel de empleo, es decir, se pasa del punto E al punto F.

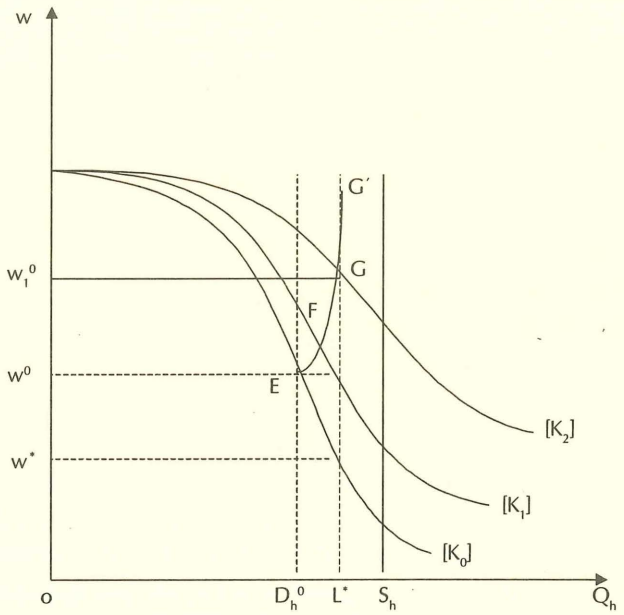
Un incremento en la oferta de trabajo no tendrá efecto alguno en los valores de equilibrio de las variables endógenas, excepto que aumenta el desempleo. El equilibrio general se da con desempleo, hecho que implica que más trabajo solamente elevará la cantidad de desempleados en el sistema.

Para el caso en el que los valores de equilibrio general inicial se dan en $u = u^*$, las relaciones de causalidad son diferentes. En este caso, la economía está en equilibrio en el nivel de empleo L^*. El empleo no puede ir más allá de este nivel, debido a que la curva de productividad del trabajo se desplazaría hacia abajo y ocasionaría que las ganancias cayeran. Entonces los capitalistas están interesados en mantener este nivel de empleo.

Considérese, en primer lugar, el efecto de un incremento en la oferta de dinero. Si la oferta de dinero aumentase, el nivel de precios aumentaría y el salario real caería por debajo de w^*, pero las firmas evitarán este resultado elevando el salario nominal. Entonces, se restaurará el equilibrio en el mismo nivel de salario real y en el mismo nivel de empleo. El dinero ahora es neutral. Entonces, la política monetaria resultará ineficaz para lograr el objetivo de pleno empleo. El nivel máximo de empleo será L^*. La tasa de desempleo u^* operará como una «tasa natural de desempleo».

Una caída en el salario nominal no nos llevará a la obtención del pleno empleo. Esta caída implicaría un nivel de em-

Figura 3.3
Efecto de la acumulación de capital sobre el equilibrio en el mercado laboral



pleo por encima de L^* y una tasa salarial por debajo de w^*. El esfuerzo de los trabajadores y su productividad disminuirían. Las firmas tendrían que incrementar el salario nominal y restaurarían L^* y w^*. Pero el salario nominal no puede bajar, por el supuesto de la teoría. Sin embargo, este resultado muestra que, aun si el salario real fuese plenamente flexible, reducir los salarios nominales no eliminaría el desempleo.

Un incremento en el capital físico genera un desplazamiento hacia arriba de la curva de demanda de trabajo. Cada firma buscaría contratar más trabajadores con el fin de obtener mayores ganancias. Si las firmas actuaran de esta manera, el empleo total tendería a elevarse por encima de nivel L^*, y eso podría

reducir la intensidad de trabajo óptima. Las ganancias totales caerían debido a que las motivaciones individuales llevan a un resultado agregado incongruente. Luego, las firmas, por separado, estarían forzadas a incrementar el salario nominal y elevar el salario real con el fin de mantener la misma ratio de desempleo y de intensidad de trabajo. Debido a que el nivel de precios cae, este efecto refuerza el alza en la tasa de salario real. Como consecuencia de esto, el producto se eleva.

Una acumulación continua de capital produce un aumento en el valor de w^*, pero L^* permanecerá constante. Entonces, la acumulación de capital no puede eliminar el desempleo. El efecto de un incremento en el stock de capital en el equilibrio del mercado de trabajo puede ser visto en la Figura 3.3. Un incremento de K_1 a K_2 en el *stock* de capital eleva la tasa salarial pero no el empleo. Este hecho se aprecia como un cambio del punto G al punto G'.

¿Qué efecto produce la acumulación de capital en la distribución? Un mayor *stock* de capital ocasiona un incremento en la masa salarial, debido a que se da un incremento en la tasa de salario real. El nivel de empleo es fijo, pero el producto total se eleva por el alza del *stock* de capital. El cambio en las ganancias es ambiguo, de ahí que el cambio en la distribución sea indeterminado.

Finalmente, considérese un incremento de la oferta de trabajo. Este no produce ningún cambio en el valor de u^*. El desempleo es una medida disciplinaria para los trabajadores basada en la tasa de desempleo, no en el nivel de desempleo. Dada la cantidad demandada de trabajo, un aumento en la oferta de trabajo implicará una tasa de desempleo más elevada. Esto es cierto y no importa si la situación de equilibrio general inicial fuese con un nivel de desempleo menor o igual que L^*. Por lo tanto, nada cambia en el equilibrio inicial, excepto que tanto la tasa de desempleo como la cantidad de desempleo absoluto se elevan. No obstante, el valor de los umbrales se modifica: el de w^* cae y el de L^* se eleva.

En el Cuadro 3.1, se presenta la matriz de las proposiciones beta que se derivan del modelo de la economía épsilon. En la primera columna, debajo de cada variable exógena, se describen las relaciones de causalidad bajo el supuesto de que el equilibrio inicial ocurre con una tasa de desempleo por encima de la tasa de desempleo de eficiencia ($u>u^*$). La segunda columna muestra las relaciones de causalidad bajo el supuesto de que el equilibrio inicial ocurre cuando $u=u^*$.

En suma, en la economía épsilon, la tasa de desempleo persistirá sin importar en que dirección o magnitud varíe una variable exógena. El valor de u^* puede ser denominado como *tasa de desempleo de eficiencia.* A esta tasa, el conjunto de ganancias factibles será mayor. El valor de w^* puede ser denominado como *tasa de salario de eficiencia.* Mientras u^* fija el límite superior en el empleo, w^* es el límite inferior del salario real consistente con u^* —dadas las condiciones del mercado de trabajo—.

Cuadro 3.1
Matriz de proposiciones beta de un modelo épsilon estático y cerrado

Variables endógenas	Variables exógenas					
	S_m		K_b		S_h	
w	−	0	+	+	0	0
D_h	+	0	+	0	0	0
U	−	0	−	0	+	+
Q_b	+	0	+	+	0	0
P	+	0	?	?	0	0
W	?	0	+	+	0	0
P/W	?	0	?	?	0	0
P_b	+	+	−	−	0	0
P_h	0	+	0	+	0	0

Nota: En la primera columna el equilibrio inicial y final se da con $u > u^*$, mientras que en la segunda columna se da con $u = u^*$.

Los cambios en las condiciones del mercado de trabajo —cantidad ofertada de trabajo y *stock* de capital— modifican el valor del salario de eficiencia w^* pero no la tasa de desempleo de eficiencia u^*. Un incremento continuo en la oferta de dinero elevará los niveles de empleo (así como el nivel del producto y de los precios) hasta el nivel de L^*; más allá de este umbral, mayor oferta de dinero elevará solo el nivel de precios. Similarmente, un incremento continuo en el *stock* de capital elevará el nivel de empleo (así como el nivel del producto y de los salarios) hasta el nivel de L^*; más allá de este umbral, la mayor acumulación de capital incrementará el producto total y el salario real.

En lo que se refiere a los cambios en la distribución de ingresos, las predicciones del modelo pueden ser resumidas como sigue: las variables exógenas del sistema tienen un efecto ambiguo o nulo en la ratio de las ganancias y la masa salarial. La matriz de las proposiciones beta presentadas en el Cuadro 3.1 muestra tales resultados.

3.6 Consistencia empírica: los países del primer mundo

¿Se parecen los países del primer mundo a la economía épsilon? Este modelo predice que el desempleo es la característica básica de estos países, predicción consistente con los datos empíricos (Barro 1997: 225).

Sin embargo, este modelo también predice que, en el corto plazo, el salario real y el producto se mueven en direcciones opuestas como resultado de los cambios en la oferta de dinero. El salario real es contracíclico. Esto parece ser inconsistente con los datos empíricos. Por ejemplo, en Estados Unidos, el salario real es procíclico. Barro reporta esto para el periodo 1959-1996: el coeficiente de correlación entre el salario real y el PBI real, una vez corregidas sus tendencias, es positivo e igual a 0.55 (Barro 1997: 224). La oferta de dinero (o gasto de gobierno financiado con crédito del Banco Central) no puede reproducir

este hecho estilizado. La raíz de esta inconsistencia es fácil de identificar. El efecto de la política monetaria (equivalente a la política fiscal en este modelo) funciona a lo largo de una curva de demanda de trabajo dada. De ahí que el modelo no pueda predecir que el salario real y el producto se muevan en la misma dirección.

¿Se parecen los países de tercer mundo a la economía épsilon?. Estos países muestran tasas de desempleo positivas, particularmente en las áreas urbanas. Pero estos países muestran una cantidad significativa de auto empleo de la fuerza laboral —auto empleo que este modelo teórico no predice—.

Con el fin de esclarecer si la teoría épsilon puede explicar el funcionamiento de los países del primer mundo se necesita construir otro modelo. Algunos factores han sido dejados de lado en el modelo presentado en este capítulo. Siguiendo el algoritmo establecido por la metodología alfa-beta, se construirá ahora un nuevo modelo en el que el supuesto de economía cerrada es sustituido por el de una economía abierta y en el que la tecnología está sujeta a factores limitantes. La estructura del modelo previo será utilizada y será muy útil en la construcción de este nuevo modelo. Este nuevo modelo se desarrollará en el próximo capítulo.

Capítulo 4
Un modelo abierto de épsilon

La economía épsilon operará ahora como una economía abierta. Considérese primero el comercio internacional de mercancías manteniendo el supuesto de que épsilon produce solamente la mercancía B. La manera más simple de introducir el comercio internacional es suponer que épsilon exportará la mercancía B e importará la mercancía C —mercancía que será usada como insumo en la producción del bien B—. Todas las importaciones se consideran como un bien intermedio. El comercio internacional necesita de un mercado de divisas —mercado en cuyo funcionamiento interviene el Gobierno—. Luego, la divisa puede ser comprada del y vendida al gobierno a una tasa de cambio dada. El mercado de divisas se supone que es un mercado walrasiano; su precio es flexible.

4.1. Un modelo con tipo de cambio flexible

Existen ahora cinco mercados en la economía: trabajo (H), mercancías (B) y (C), dinero doméstico (M) y divisas (E). Se asume que esta economía es muy pequeña en el ámbito del comercio internacional, de modo que los precios internacionales están dados exógenamente. Los bienes B y C son ambos transables. Por lo tanto, el precio doméstico de las mercancías B y C se forma de la siguiente manera:

$$P_b = P_e P_b^* \qquad (4.1)$$
$$P_c = P_e P_c^*$$

Aquí, P_e es el tipo de cambio nominal y P_b^* y P_c^* son precios internacionales denominados en unidades de moneda extranjera. El precio doméstico de cada bien debe ser igual al precio internacional multiplicado por el tipo de cambio nominal. Si esta igualdad no se mantiene, las personas comprarían la mercancía B de la fuente más barata y la venderían donde el precio fuese mayor. La posibilidad de este arbitraje asegura la formación de los precios indicada. La ratio $z^* = P_b^* / P_c^*$ se define como los términos de intercambio internacional.

Las firmas domésticas producen el bien B bajo la lógica de maximización de beneficios nominales. Para la firma *j*, esta racionalidad puede ser presentada como sigue:

$$\text{Maximiza} \quad P_j' = P_b Q_{bj} - P_h D_{hj} - P_c D_{cj}, \tag{4.2}$$

sujeto a la restricción de la función de producción (donde $\lambda^* = 1$):

$$Q_{bj} = F_j (D_{hj}, K_{bj})$$
$$Q_{bj} = D_{cj} / n$$
$$X_{bj} = (1/z^*) D_{cj}$$

Aquí, *n* es un coeficiente tecnológico, las unidades del bien C requeridas para producir una unidad del bien B. La función de producción está sujeta a *factores limitacionales* y es, por lo tanto, representada como un sistema de ecuaciones.[1] La última ecuación indica que la firma debe pagar con parte de su producto por las cantidades del bien importado D_c.

La condición de equilibrio de la firma es la siguiente:

$$F_{j1} (D_{hj}, K_{bj}) (P_b - n P_c) = P_h \tag{4.3}$$
$$F_{j1} (D_{hj}, K_{bj}) (1 - n / z^*) = w, \text{ donde } (n / z^*) < 1, \text{ y } w \geq w^*$$

[1] Un *factor limitacional* se define como sigue: un aumento en su cantidad es una condición necesaria pero no suficiente para un aumento en el nivel del producto. Esta definición implica que este factor de producción no puede ser sustituido por otros. En la tecnología aquí asumida, el capital y el trabajo pueden ser sustitutos entre sí, pero no existe sustitución entre el insumo C y el factor combinado capital-trabajo.

La productividad marginal del trabajo es ahora neta del costo de utilización del insumo limitacional C. La primera ecuación muestra la curva de demanda de trabajo en términos de la tasa de salario nominal (P_h) y la segunda en términos de la tasa de salario real (w). Por lo tanto, la cantidad de trabajo demandada depende también de los términos de intercambio.[2]

Dado que la situación de equilibrio indicada en (4.3) es estable y única, la estática comparativa puede ser aplicada para determinar proposiciones beta para la firma. Luego, la función de demanda laboral de la firma puede derivarse de (4.3); dicha función indicará que el precio del bien B y el *stock* de capital tienen un efecto positivo, mientras que el precio del bien C y el salario nominal tienen un efecto negativo. La cantidad ofrecida del bien B es obtenida de la restricción presupuestaria de la firma, derivada de (4.2).

Agregando sobre todas las firmas, estas funciones pueden ser escritas de la siguiente forma:

$$\begin{aligned} &D_h = H\,(P_b, P_c, P_h, K_b),\ H_1>0,\ H_2<0,\ H_3<0,\ H_4>0 \\ &P_b S_b = P_b Q_b - P_b X_b - P' = (P_c D_c - P_b X_b) + P_h D_h \\ &X_b = (1/z)\, D_c = f\,(D_h, K_b) \end{aligned} \qquad (4.4)$$

La primera ecuación es la función de demanda laboral. La segunda ecuación muestra la cantidad ofrecida del bien B al mercado doméstico; dicha ecuación también representa las ecuaciones presupuestarias de la firma. La tercera ecuación señala que las firmas deben pagar por sus importaciones mediante la exportación de parte de su producción.

[2] Dimensionalmente, la ratio n/z^* es un número puro. La condición de que $(n/z^*)<1$ asegura que el sistema tecnológico es productivo. Si una firma intercambia en el mercado una unidad del bien B por dos unidades del bien C como insumo —que ayudan a producir menos de una unidad—, lo único que ha hecho es reducir su *stock* del bien B. Para que el sistema sea productivo, estas dos unidades del bien C deben ayudar a producir más de una unidad del bien B.

¿Cómo se logra el equilibrio en el mercado de divisas? Bajo un sistema de tipo de cambio flexible, se puede suponer un mecanismo simple en el funcionamiento del mercado de divisas. Las firmas exportan la mercancía B y obtienen divisas que luego venden al banco central para obtener moneda doméstica. Luego, las firmas utilizan esa moneda doméstica para comprar divisas del banco central, y con estas divisas pagan por las importaciones de la mercancía C. Las transacciones se hacen a precios de mercado dados. Los costos de transacción de estas operaciones son cero.[3]

En este modelo particular, el equilibrio en el mercado de divisas es determinado por el comportamiento de las firmas. Las firmas al pagar por sus importaciones exportando sus bienes a un tipo de cambio dado aseguran que el mercado de divisas esté en equilibrio. Este intercambio es parte de la restricción presupuestaria de las firmas, pues las importaciones se refieren a bienes intermedios solamente. Dado el *stock* de capital, el nivel de empleo determinará la cantidad producida del bien B, pero esta producción no será factible a menos que se importen las cantidades del insumo C que son tecnológicamente requeridas, y por estas importaciones las firmas pagan exportando parte de esa producción. Esta relación es la que muestra la tercera ecuación del sistema (4.4).

De las restricciones presupuestarias de las firmas, las siguientes relaciones son ciertas:

$$P_c D_c = P_b X_b \qquad (4.5)$$

$$P_c^* D_c = P_b^* X_b \qquad (4.6)$$

[3] Si esta economía produjese varios bienes, algunas firmas exportarían bienes mientras que otras los importarían. El valor de las exportaciones debería ser igual al valor de las importaciones si este mercado es walrasiano. En un mundo de un único bien, como es el caso del presente modelo, el comercio internacional opera *como si* la misma firma tuviese que exportar para pagar sus importaciones.

La primera ecuación muestra el intercambio de las firmas: pagan con exportaciones las importaciones requeridas. La segunda ecuación indica el equilibrio correspondiente en el mercado de divisas. Este equilibrio es independiente del valor del tipo de cambio P_e.

La curva de demanda laboral en términos de la tasa de salario real se presenta en la Figura 4.1, panel (c). La curva M'R' representa la productividad marginal bruta del trabajo, donde el costo para la firma de obtener el necesario insumo C se deduce del producto total. Para un nivel dado de empleo, la brecha entre la productividad marginal del trabajo y la contraparte neta la determina el valor de los términos de intercambio. Luego, para una tasa de salario real dada, las firmas emplean trabajadores en la cantidad indicada por la productividad marginal neta del trabajo. En este nivel de empleo, existe equilibrio en el mercado de divisas.

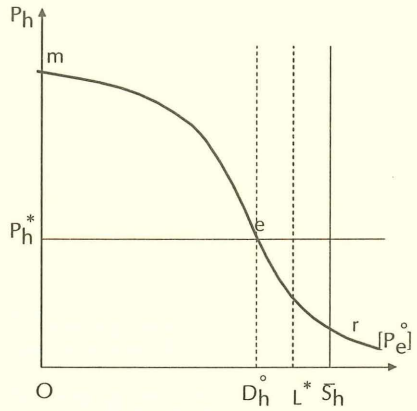
Debido a que la intensidad del trabajo depende de la tasa de desempleo, existe una tasa de desempleo mínima que asegura dicha intensidad (dada la oferta de trabajo, el nivel máximo de empleo que es consistente con el grado de esfuerzo requerido por las firmas es L^*). Por lo tanto, el segmento MN en la Figura 4.1 (c) representa la curva de demanda de trabajo. Además del *stock* de capital, la curva de demanda tiene ahora un nuevo parámetro: el coeficiente de términos de intercambio. Luego, una caída en este coeficiente causa un desplazamiento hacia abajo en la curva de demanda de trabajo, y un aumento desplaza la curva hacia arriba.

Los trabajadores ganan salarios en moneda doméstica y consumen solamente el bien B. Su comportamiento será determinado por las mismas variables que se presentaron más arriba. En una economía abierta, las funciones de demanda agregada del bien B y del dinero pueden ser escritas como sigue:

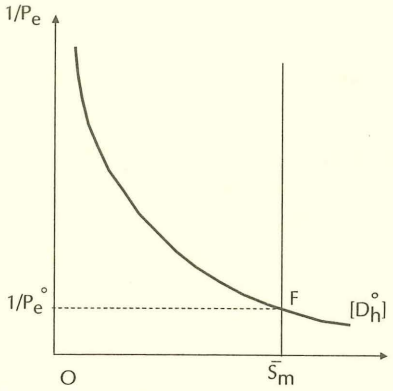
$$\begin{aligned} &D_b = F(P_b, D_h, P_h), \text{ donde } F_1<0,\ F_2>0,\ F_3>0 \\ &D_m = P_b\, L(D_b), \text{ donde } L'>0,\ L''<0 \\ &\text{Sujeto a } P_h D_h = P_b D_b + (D_m - S_m) \end{aligned} \qquad (4.7)$$

Figura 4.1
Equilibrio general en la economía épsilon

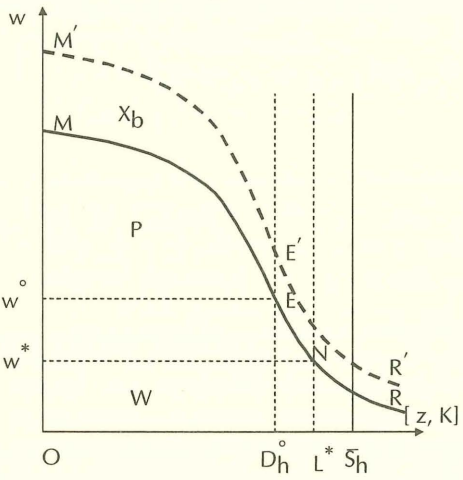
Mercado laboral
(a)



Mercado del dinero
(b)



(c)



4.2 Equilibrio general

Las condiciones de equilibrio general pueden ahora ser establecidas. Existen tres mercados walrasianos y uno no-walrasiano (el de trabajo). Estas condiciones son:

$$\begin{array}{lll} \text{Mercado laboral} & : & S_h = D_h + U \\ \text{Mercado del bien B} & : & S_b = D_b \\ \text{Mercado monetario} & : & S_m = D_m \\ \text{Mercado de divisas} & : & P_b^* X_b = P_c^* D_c, \end{array} \qquad (4.8)$$

sujeto a la restricción del presupuesto agregado:

$$P_b (D_b - S_b) + P_e (P_c^* D_c - P_b^* X_b) + (D_m - S_m) = 0$$

El equilibrio general se resuelve de manera secuencial. Las tres primeras ecuaciones constituyen un subsistema, el mismo que se resuelve simultáneamente. El mercado de divisas no afecta este subsistema; más bien, es afectado por la solución del subsistema. Debido a que el subsistema contiene dos mercados walrasianos, uno de ellos es redundante por la ley de Walras. Sea el mercado del bien B el redundante. Luego, solo hay que resolver simultáneamente en los mercados de trabajo y de dinero.

El equilibrio del mercado de dinero puede establecerse ahora en términos del tipo de cambio, en vez del nivel del precio. A mayor tipo de cambio, mayor el precio doméstico y, por lo tanto, mayor la demanda de dinero nominal. Una curva de demanda de dinero de pendiente negativa puede obtenerse utilizando 1/Pe como el «precio del dinero». Dado un *stock* de dinero, habrá un valor de tipo de cambio que limpie el mercado de dinero.

En términos del análisis de equilibrio parcial, el tipo de cambio está determinado por el mercado de dinero. Si aumenta la oferta monetaria, las personas querrán reducir sus nuevas tenencias y comprar más cantidades del bien B, y eso reduciría las exportaciones. La oferta de divisas cae y, por lo tanto, el tipo de cambio aumenta para limpiar el mercado de dinero. El equilibrio en el mercado de divisas se restaura también. Si la deman-

da de dinero aumenta debido a un aumento en el nivel de empleo, las personas buscarán acumular saldos reales y deberán reducir sus compras del bien B, y eso liberaría cantidades de este bien para ser exportadas. La cantidad ofrecida de divisas aumenta y, para limpiar el exceso de oferta, el tipo de cambio debe caer.

El equilibrio general puede, por lo tanto, ser determinado resolviendo dos ecuaciones, aquellas que representan los equilibrios respectivos en el mercado de dinero y en el mercado laboral. De (4.1), (4.4) y (4.7), pueden ser escritas como sigue:

$$D_h = H\,(P_e; P_h, P_b^*, P_c^*, K_b),\ H_1>0,\ H_2<0,\ H_3>0,\ H_4<0,\ H_5>0$$
$$S_m = P_e P_b^*\, L\,(D_b) = M\,(P_e, D_h; P_h, P_b^*),\ M_i>0, \qquad (4.9)$$

sujeto a $u \geq u^*$ y $w \geq w^*$

Estas dos ecuaciones deben resolver, para dos variables endógenas, el tipo de cambio P_e y el nivel de empleo D_h. La solución de este sistema es simultánea. Una vez que se conocen, el resto de variables endógenas será también conocido.

Hay que notar que en este sistema existe un grado de libertad para escoger la variable exógena. Si la oferta monetaria fuese exógena, el tipo de cambio nominal sería endógeno; si el tipo de cambio nominal fuese exógeno, la oferta monetaria sería endógena. El primer caso da lugar a un régimen de tipo de cambio flexible, mientras que el segundo da lugar a un régimen de tipo de cambio fijo. En este estudio, la oferta monetaria continuará siendo exógena.

Una vez que se ha resuelto el subsistema, se puede resolver por el equilibrio en el mercado de divisas. Conocido el valor del nivel del empleo, el producto total queda determinado y también la cantidad del insumo importado D_c. Luego, la cantidad exportada del bien B queda también determinada X_b, dado que las firmas tienen que pagar por sus importaciones a los términos de intercambio dados. Así, el equilibrio en el mercado de divisas queda establecido.

Pero ahora el nivel del ingreso nacional no es igual al producto total, pues parte del producto ha sido utilizado para pagar por los insumos importados. El nivel del ingreso nacional es, por lo tanto, igual al producto total neto —neto luego de pagar con parte del producto por las importaciones del insumo utilizado en su producción—. Este nivel de producto neto es el que se distribuye como salarios y ganancias. Una vez conocido el tipo de cambio de equilibrio, el nivel de precios queda también determinado. Este, a su vez, determina el salario real. La masa salarial real queda entonces conocida. La ganancia es igual al producto total menos la cantidad exportada y menos la masa salarial.

La Figura 4.1 muestra la situación de equilibrio general. El panel (a) muestra el mercado de dinero en términos de las tasas de salario nominal —donde el tipo de cambio es un parámetro de la curva de demanda de trabajo—. El panel (b) presenta el mercado de dinero en términos de los tipos de cambio —donde el nivel de empleo es un parámetro de la curva de demanda de dinero—. Existe un único conjunto de valores de equilibrio de tipo de cambio y nivel de empleo.

El panel (c) representa la curva de demanda de trabajo en términos de tasas de salario real. Conocido el tipo de cambio de equilibrio, el nivel de precios queda determinado y también el salario real. Este salario real determina el nivel de empleo que maximiza la ganancia capitalista y que coincide con el valor de equilibrio ya determinado. En esta figura, el conjunto de valores del producto total bruto que se puede obtener con cada nivel de empleo está representado por la curva superior M'E'R'. La curva inferior representa el conjunto de valores del producto neto correspondiente, pues se resta la cantidad de exportaciones que son necesarias para importar los insumos necesarios para obtener ese nivel de producto total, dado el coeficiente de los términos de intercambio. Esto significa que cualquier nivel de empleo —que dará lugar a una cantidad de producto neto en la curva MER— tiene asegurado un equilibrio en el mercado de divisas.

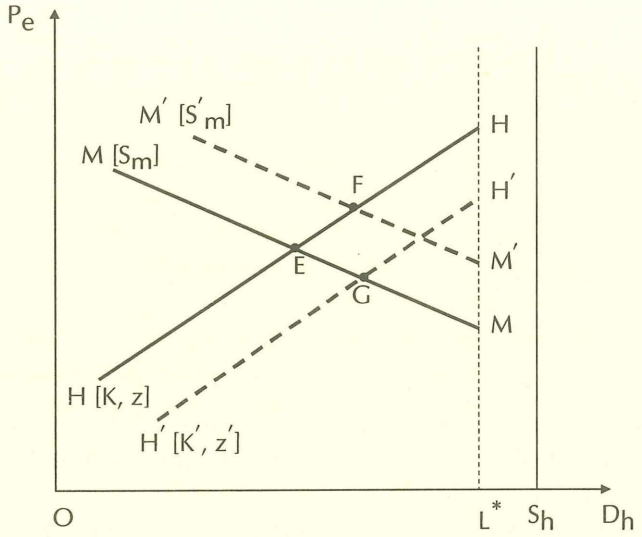
Las firmas, consistentes con su lógica de la búsqueda de la máxima ganancia, igualan el salario real a la productividad marginal neta del trabajo, y así se determina el nivel de empleo. Ciertamente, el área bajo la curva M'E' no representa el ingreso nacional. El ingreso nacional es igual al área debajo de la curva ME. El producto bruto es mayor, pues una parte (X_b) se destina a exportaciones, de modo que se puedan importar las cantidades del insumo C requeridas para producir el bien B, dado los términos de intercambio. Luego, este gráfico muestra que, en efecto, el mercado de divisas está en equilibrio.

También esta figura muestra que el ingreso nacional se distribuye entre beneficios y salarios. Este es el equilibrio en la distribución del ingreso. Esta figura también permite mostrar que existe equilibrio en el mercado del bien B. Para ello es suficiente percatarse de que la cantidad ofrecida al mercado doméstico, que es igual al producto total neto menos los beneficios, es un área que coincide con el área de la masa salarial, que constituye la cantidad demandada domésticamente.

4.3. Proposiciones beta

Como se mostró arriba, es suficiente resolver las interacciones entre dos mercados, trabajo y dinero, para determinar el tipo de cambio y el nivel de empleo, y obtener así el equilibrio general de la economía épsilon. Un simple gráfico —una consolidación de los paneles (a) y (b) de la Figura 4.1— puede mostrar esta solución, como puede verse en la Figura 4.2. La curva HH, de pendiente positiva, representa el conjunto de combinaciones de tipo de cambio y nivel de empleo consistentes con el equilibrio en el mercado de trabajo. A mayor tipo de cambio, mayor el nivel de precio doméstico y mayor el nivel de empleo. La curva con pendiente negativa, MM, representa el conjunto de combinaciones de tipo de cambio y nivel de empleo que limpia el mercado de dinero. A mayor nivel de empleo, mayores los

Figura 4.2
Equilibrio en los mercados monetario y laboral



saldos reales requeridos, menor la cantidad doméstica demandada del bien B, mayor la cantidad ofrecida de divisas y menor el tipo de cambio. El conjunto de valores de tipo de cambio y nivel de empleo consistente con el equilibrio en ambos mercados se determina por la intersección de ambas curvas —como en el punto E—.

Este equilibrio es claramente estable. Por lo tanto, se puede aplicar la estática comparativa para derivar proposiciones beta. Las variables exógenas del sistema incluyen la cantidad de dinero (un parámetro en la curva MN), el *stock* de capital y los términos de intercambio (parámetros en la curva HH), y la oferta de trabajo. Además, el sistema opera con un máximo nivel de empleo L^*. Primero, considérese el caso en el que la situación de equilibrio inicial ocurre en la situación $u>u^*$, es decir, por debajo del empleo L^*.

El efecto de un cambio en la oferta de dinero puede ser determinado formando un sistema de dos ecuaciones con las deri-

vadas parciales respecto a la oferta monetaria del sistema (4.9). Ellas son las siguientes:

$$\begin{aligned} \partial D_h/\partial S_m - H_1\ \partial P_e/\partial S_m &= 0 \\ M_2\ \partial D_h/\partial S_m + M_1\ \partial P_e/\partial S_m &= 1 \end{aligned} \qquad (4.10)$$

La solución de este sistema implica $\partial D_h/\partial S_m > 0$ y $\partial P_e/\partial S_m > 0$. El efecto de un aumento de la oferta de dinero es positivo sobre el tipo de cambio y el nivel de empleo. Una mayor oferta de dinero resultará en un exceso de saldos reales, por lo que la demanda doméstica por el bien B se elevará, la oferta de divisas caerá y el tipo de cambio aumentará. Una nueva situación de equilibrio aparecerá. En la Figura 4.2, la curva MN se desplaza hacia arriba y el nuevo equilibrio sería el punto F, hecho que implica un mayor nivel de tipo de cambio y un mayor nivel de empleo.

El efecto de un aumento en el precio internacional del bien B, exportado por la economía épsilon —un aumento en los términos de intercambio—, puede ahora ser determinado resolviendo el sistema siguiente:

$$\begin{aligned} \partial D_h/\partial P_b - H_1\ \partial P_e/\partial P_b &= H_3 \\ M_2\ \partial D_h/\partial P_b + M_1\ \partial P_e/\partial P_b &= -M_3 \end{aligned} \qquad (4.11)$$

La solución de este sistema implica $\partial D_h/\partial P_b^* > 0$, mientras que $\partial P_e/\partial P_b^* < 0$. Un aumento en los términos de intercambio tendrá el efecto de elevar el nivel de empleo y reducir el tipo de cambio. Un aumento en el precio del bien B aumentará la demanda laboral, que a su vez aumentará la demanda de dinero y reducirá el tipo de cambio. Una nueva situación de equilibrio surgirá. En la Figura 4.2, un aumento en los términos de intercambio desplaza la curva HH hacia fuera. El nuevo equilibro está en el punto G, hecho que implica un mayor nivel de empleo y un tipo de cambio menor.

Puede mostrarse que un aumento en el precio internacional del bien C causa un efecto contrario, una caída en el nivel de empleo y un aumento en el tipo de cambio. Queda claro que la función que muestra la productividad marginal del trabajo es homogénea de grado cero respecto al precio internacional de los bienes B y C. Por lo tanto, el efecto solamente tiene que ver con la ratio, con los términos de intercambio.

En términos de la distribución del ingreso, el nuevo equilibrio obviamente implica un aumento en la masa salarial. El efecto en los beneficios totales es menos obvio. Si el empleo no cambiara y las firmas pagasen a los trabajadores su nueva productividad marginal de trabajo, los beneficios aumentarían. La razón es que el desplazamiento en la curva de demanda de trabajo no es paralelo; es decir, la nueva curva es más empinada. Sin embargo, el nivel de empleo de equilibrio aumenta, lo que claramente implica que los beneficios aumentarán. Luego, la ratio de beneficios a masa salarial es indeterminada.

El efecto de la acumulación de capital puede ser determinado resolviendo el siguiente sistema:

$$\begin{aligned} \partial D_h/\partial K \; - H_1 \; \partial P_e/\partial K &= H_5 \\ M_2 \; \partial D_h/\partial K \; + M_1 \; \partial P_e/\partial K &= 0 \end{aligned} \qquad (4.12)$$

La solución de este sistema implica: $\partial D_h/\partial K > 0$ y $\partial P_e/\partial K < 0$. El efecto de un incremento en el *stock* de capital es positivo sobre el valor de empleo de equilibrio y negativo sobre el valor de equilibrio del tipo de cambio. Estos resultados son similares a aquellos obtenidos en el caso de un aumento en los términos de intercambio. En la Figura 4.2, la curva HH se desplaza hacia fuera, esta vez como resultado del aumento del *stock* de capital, y el nuevo equilibrio está en el punto G. Este resultado implica que un aumento en los términos de intercambio es equivalente a una destrucción del *stock* de capital. Sin embargo, el efecto de los términos de intercambio debe ser visto como un efecto de corto plazo, y el de la acumulación de capital como uno de largo plazo.

La matriz de proposiciones beta se muestra en el Cuadro 4.1. Bajo cada variable exógena, la primera columna se refiere a las relaciones causales cuando la situación de equilibrio inicial se da con un desempleo mayor a u^*. Estos son los efectos determinados anteriormente. La segunda columna corresponde al caso en el que en la situación de equilibrio general se tenía $u=u^*$. Estos resultados pueden derivarse de la Figura 4.2, una vez que se reconoce que las curvas HH y MM son válidas solo para rangos menores a L^*.

En el nivel de empleo L^*, la economía opera como si estuviese a «pleno empleo», de manera que los efectos de cambios en la oferta monetaria son solo sobre precios. El tipo de cambio, el nivel de precios domésticos y la tasa de salario nominal aumentan todos en la misma proporción. El salario nominal aumenta porque, de otro modo, el salario real estaría por debajo del valor de salario de eficiencia w^*, y el grado de intensidad de

Cuadro 4.1
Matriz de proposiciones beta de un modelo épsilon estático y abierto

Variables endógenas	Variables exógenas					
	S_m		z^*		K_b	
P_e	+	+	−	−	−	−
D_h	+	0	+	0	+	0
U	−	0	−	0	−	0
w	−	0	+	+	+	+
Q_b	+	0	+	+	+	+
W	?	0	+	+	+	+
P	+	0	+	+	?	?
P/W	?	0	?	?	?	?
P_b	+	+	−	−	−	−
P_h	0	+	0	+	0	+

Nota: La primera columna se refiere a situaciones donde $u > u^*$; y la segunda a aquellas donde $u = u^*$.

trabajo caería. Un aumento en el *stock* de capital aumentaría la demanda de trabajo. El salario real no podría permanecer constante porque ello implicaría un nivel de empleo más allá de L^*, y eso ha ha sido descartado por un supuesto. En la nueva situación de equilibrio, el salario real aumenta. Un aumento en los términos de intercambio tiene el mismo efecto.

El conjunto de variables exógenas del sistema de equilibrio general incluye ahora los términos de intercambio. Un aumento en los términos de intercambio desplaza la productividad marginal del trabajo hacia arriba. Luego, un aumento en los términos de intercambio originará un mayor nivel de empleo, producto y salario real; una caída en los términos de intercambio causará el efecto contrario. Luego, en el corto plazo, este modelo predice que los *shocks* externos de precios generarán movimientos en la misma dirección para el salario real, empleo y producto. Esta asociación positiva entre variables endógenas es ahora consistente con los hechos observados en el primer mundo, hecho que no ocurría en el caso del modelo de una economía cerrada.[4]

En resumen, en este modelo de una economía épsilon abierta, los precios y cantidades son determinados simultáneamente. Esta solución implica una distribución particular del ingreso. Luego, la producción y la distribución se determinan simultáneamente. Estos resultados han supuesto que el mercado laboral sea no-walrasiano. Otro mercado que le otorgará particularidades al funcionamiento de la economía épsilon es el de crédito, pues se supondrá que también opera como un mercado no-walrasiano.

[4] El mismo efecto puede ser obtenido con una caída en la tarifa sobre las importaciones, que se utiliza como un instrumento de politica doméstica. Esto reducirá el precio doméstico de la mercancía C y desplazará la curva de demanda de trabajo neto hacía arriba.

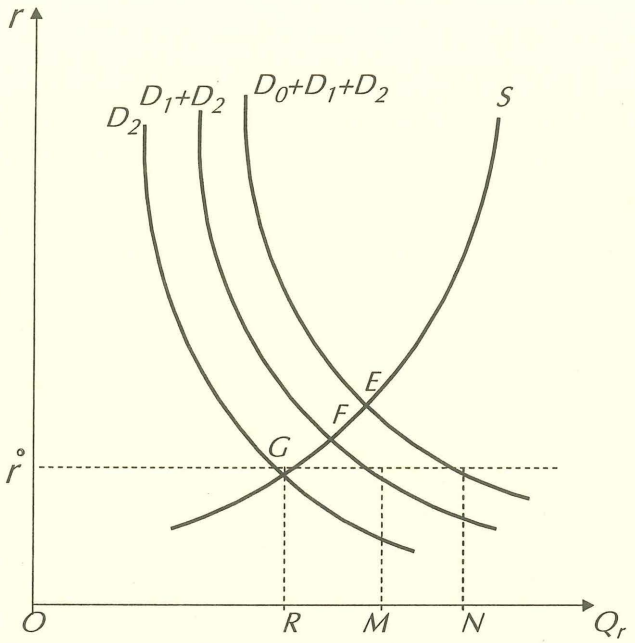
4.4. Un modelo con mercado de crédito bancario

Suponga ahora la existencia de un mercado de crédito bancario en el que las firmas puedan pedir dinero prestado. Normalmente se supone que el mercado de crédito opera bajo un contexto de información asimétrica. Las firmas están mejor informadas que los bancos respecto a su real valoración crediticia. Los incentivos de ambas partes son incompatibles, y el incumplimiento en el repago es posible. Los bancos y prestatarios tienen un problema del tipo principal-agente, similar al de las relaciones entre firmas y trabajadores estudiados anteriormente. Luego, los bancos enfrentan problemas de selección adversa y riesgo moral. Todos estos supuestos son estándar en la literatura financiera.

Algunos supuestos adicionales serán introducidos ahora. Defínase la demanda por crédito como la cantidad de dinero que los individuos están dispuestos a pedir prestado y pagar a cada tasa de interés de mercado. Esta curva de demanda tiene pendiente negativa. En cada punto de la curva de demanda, los prestatarios solo toman aquellos riesgos que son tolerables dada sus dotaciones de capital (como se muestra en el apéndice). Supóngase que la cantidad ofrecida de fondos prestables tenga pendiente positiva. Luego, habrá una tasa de interés de mercado que limpia cualquier exceso de demanda u oferta. El mercado de créditos sería walrasiano. Esto se muestra en la Figura 4.3 en el punto E.

Pero, ¿puede el mercado de crédito operar de este modo? ¿Cuáles serían los incentivos económicos de los prestatarios para devolver el préstamo? Ninguno. Si el individuo siempre puede obtener un crédito, independientemente de su historial de pagos, ¿por qué debería molestarse en devolver el préstamo? Los prestatarios no pagarían ningún costo si incumplen en devolver. La tasa de repago sería muy baja y los bancos no podrían subsistir. El mercado de crédito no puede operar como un mercado walrasiano.

Figura 4.3
Equilibrio en el mercado de crédito



Los bancos buscarán mecanismos para asegurar la devolución del préstamo. Un mecanismo posible sería fijar la tasa de interés por debajo del precio walrasiano. Se generaría así un exceso de demanda y el mercado de crédito operaría con racionamiento de cantidades. El costo para el prestatario de incumplir en la devolución sería el pasar a estar excluido del mercado. En una economía estática, sin embargo, el prestatario podría hacerse de un capital tomando un crédito por única vez y sin devolverlo. Luego, este mecanismo no operaría adecuadamente.

Otro mecanismo sería exigirle al prestatario un colateral. Supóngase que el colateral es equivalente al monto del préstamo y que, en caso de incumplimiento, el colateral puede ser eje-

cutado a costo cero. Luego, el préstamo está completamente asegurado. Pero, ahora, algunos individuos, deseosos de pedir prestado y capaces de reembolsar el préstamo, estarán excluidos del mercado. Este será el caso de aquellos que no poseen capital. Este mecanismo origina una caída en la demanda de crédito para algunos prestatarios potenciales y, por lo tanto, un desplazamiento hacia abajo de la curva de demanda de crédito. En la Figura 4.3, la solución de mercado ocurre ahora en el punto F.

Supóngase ahora que el costo *por prestatario* al otorgar un préstamo es constante; es decir, es independiente del tamaño del préstamo. El costo de evaluación, monitoreo, desembolso y cobro del préstamo no depende del tamaño del préstamo. El costo de transacción por prestatario es un costo fijo. Este supuesto implica que el costo unitario *por dólar* prestado decae con el tamaño del préstamo. Si el costo total de prestar $100 000 fuese el mismo que el de prestar $1000, el banco preferiría otorgar un préstamo del primer tipo en vez de 100 préstamos del segundo tipo. Debido a estas economías de escala, el conjunto factible de beneficios para el banco será mayor con préstamos más grandes que con préstamos pequeños. Luego, los bancos buscan la maximización de beneficios fijando un umbral para el tamaño del préstamo y el valor correspondiente del colateral.

Sea K_b^* el umbral de colateral. La lógica de maximización de beneficios de los bancos deja de lado a prestatarios potenciales que sí poseen capital pero en una cantidad menor al del umbral. Esta lógica origina otro desplazamiento hacia abajo en la curva de demanda de crédito —a D_2 en la Figura 4.3—. La curva D_2 es la demanda efectiva de crédito. La solución de mercado cae ahora en el punto G, donde la tasa de interés de equilibrio es igual a r^0 (se llamará «bien R» al crédito).

En términos de la dotación de capital de los individuos, por lo tanto, existen tres estratos en este modelo: aquellos que demandan crédito pero no poseen ningún tipo de capital, y que pueden ser representados mediante la curva de demanda D_0; aquellos que poseen capital pero en una cantidad menor a K_b^*,

y que pueden ser representados por la curva de demanda D_1; y aquellos con capital igual o mayor a K_b^*, y que pueden ser representados por la curva de demanda D_2.

¿Es el mercado de crédito walrasiano? No, no lo es. El mercado se «limpia» solamente para aquellos que son elegibles, los incluidos en la curva de demanda D_2. Al precio r^o existe un exceso de demanda. A este precio, hay individuos dispuestos a pedir prestado y reembolsar el préstamo, pero que no son elegibles para el crédito. El exceso de demanda es equivalente al segmento RN de la Figura 4.3. Incluye a aquellos que pueden pedir prestado y reembolsar el préstamo bajo la condición de no colateral (MN) y a aquellos que desean préstamos más pequeños (segmento RM). No todo aquel que desee intercambiar ciertas cantidades de crédito al precio de mercado prevaleciente será capaz de hacerlo. El mercado de crédito es un mercado no-walrasiano.

En resumen, esta teoría nos dice que, para maximizar beneficios, los bancos deben excluir a aquellos prestamistas potenciales dotados pobremente de bienes de capital. *Existe exclusión, y no racionamiento, en el mercado de crédito.*

Esta proposición teórica es contraria a lo que la mayoría de las teorías del mercado de crédito dicen. Una de estas teorías supone que la tasa de devolución depende negativamente de la tasa de interés porque los prestamistas se involucrarán en proyectos con altos retornos para poder reembolsar préstamos más caros, siendo dichos proyectos más riesgosos. Este supuesto implica que las tasas de reembolso caerán a medida que nos movamos hacia arriba sobre la curva de demanda de crédito. Luego, habrá una tasa de interés óptima que maximiza los beneficios esperados. A este precio, podría haber un exceso de demanda; el mercado de crédito puede ser no-walrasiano (Stiglitz y Weiss 1981). Pero queda claro que el exceso de demanda no es una solución necesaria; el equilibrio puede darse también con un exceso de oferta de crédito. Todo depende de si la tasa de interés óptima está por debajo o por encima del precio walrasiano.

Otras teorías incluyen aversión al riesgo en el comportamiento de los prestatarios. En el caso de algunos requerimientos de colateral, algunos individuos pueden decidir no pedir prestado aun cuando tuviesen el capital necesario (Blinder 1987). Luego, no todo el segmento RN representaría exceso de demanda.

¿Cuál es la forma que adopta la función de demanda de trabajo cuando las firmas necesitan crédito bancario en su proceso de producción? En el corto plazo, cuando el *stock* de capital se encuentra ya instalado, se supone que las firmas requieren crédito para obtener capital de trabajo. Deben pagar a sus factores de producción antes de que reciban ingresos por ventas. Las firmas buscan obtener crédito de corto plazo de los bancos.

Un modelo simple de la economía épsilon donde el capital de trabajo es un factor de producción limitacional se desarrollará ahora. Las firmas buscarán maximizar sus beneficios de la siguiente manera:

$$\text{Maximizar } P_j' = P_b Q_{bj} - P_h D_{hj} - P_c D_{cj} - r D_{rj}, \qquad (4.13)$$

sujeto a:

$$\begin{aligned} Q_{bj} &= F_j(D_{hj}, K_{bj}) \\ &= D_{cj}/n \\ &= D_{rj}/q P_b \\ K_{bj} &> K_b^* \end{aligned}$$

El coeficiente $q>0$ se refiere al requerimiento tecnológico del capital de trabajo en la producción del bien B. Los símbolos D_r y r representan la cantidad demandada, y la tasa de interés de mercado, respectivamente; K_b^* es el mínimo colateral para poder ser elegido en el mercado de crédito. En este mundo de un solo bien, el *stock* de capital es también la riqueza de las firmas y, por ello, es esta riqueza lo que los bancos exigen como colateral.[5]

[5] Cabe hacer una nota acerca de la dimensionalidad de la primer ecuación del sistema (4.13). Debido a que el beneficio es un flujo, suponga que su dimen-

La condición de equilibrio de la firma puede ser expresada por las siguientes ecuaciones:

$$F_{j1}(D_{hj}, K_{bj})(P_b - n P_c - q P_b r) = P_h \qquad (4.14)$$
$$F_{j1}(D_{hj}, K_{bj})(1 - n/z - q r) = w$$

La primera ecuación muestra la curva de demanda de trabajo de la firma en términos de salarios nominales y la segunda en términos de salarios reales. La productividad marginal del trabajo es neta del costo de utilizar los dos factores limitacionales, insumos importados y crédito. Hay dos desplazamientos en la curva de productividad marginal bruta del trabajo, uno por cada factor limitacional. Queda claro que la posición de la demanda de trabajo depende de los términos de intercambio z y, también, de la tasa de interés nominal r. Un aumento en la tasa de interés desplaza la curva de demanda de trabajo hacia abajo. Si la tasa de interés aumenta, la productividad marginal neta del trajo caerá, dado que el modelo supone que el crédito no puede sustituir al trabajo.

Agregando sobre todas las firmas, se obtiene la siguiente curva de demanda:

$$D_h = H(P_e, P_b^*, P_c^*, r, P_h, K_b),$$
$$\text{donde } H_1>0,\ H_2>0,\ H_3<0,\ H_4<0,\ H_5<0,\ H_6>0 \qquad (4.15)$$

El efecto de la tasa de interés sobre la cantidad de trabajo demandada es negativo. Los efectos de las otras variables son idénticos a los hallados en la función mostrada más arriba en el sistema (4.9). Esta ecuación implica que, ante una caída en la tasa de interés, la curva de demanda de trabajo en términos de salarios reales se desplazará hacia arriba.

---

sión es, digamos, «dólares/año». La tasa de interés tiene una dimensión equivalente a «1/año» y, dado que el crédito es un *stock*, tiene una dimensión de «dólar». Luego, la dimensión de r D es «dólares/año», igual que la de los beneficios.

La demanda de crédito, por ser una demanda derivada, se obtiene como sigue:

$$D_r = q\, P_b\, F(D_h, K_b) \qquad (4.16)$$

Esta ecuación dice que, a mayor nivel de empleo, mayor la cantidad demandada de crédito. Luego, todas las variables que determinan la función de demanda de trabajo también afectan —en la misma dirección— a la cantidad de crédito demandada.

En una economía abierta, los bancos pueden utilizar sus fondos prestables para comprar bonos en el exterior. Hay una decisión de portafolio involucrada. Los préstamos domésticos y los bonos externos compiten en el portafolio de los bancos con retornos dados por la tasa de interés doméstica (r) y la tasa de interés internacional (r^*). A mayor ratio entre estas tasas, mayor la oferta de fondos prestables domésticos, dada la base monetaria y los parámetros del sistema bancario —tales como la tasa de encaje legal—. El *stock* de la base monetaria (S_m) determina la demanda por depósitos bancarios, los mismos que, a su vez, determinan los fondos prestables de los bancos. Luego, la oferta de crédito puede ser representada por la siguiente ecuación:

$$S_r = N(r, r^*, S_m),\ N_1 > 0,\ N_2 < 0,\ N_3 > 0 \qquad (4.17)$$

En una economía con bancos, el concepto de dinero se refiere a circulante y demanda por depósitos. La oferta de este dinero depende del *stock* de la base monetaria y del multiplicador bancario, que puede ser tomado como dado. Luego, la oferta monetaria es aún exógena en el mercado de dinero. Bajo un análisis de equilibrio parcial, los precios y las cantidades de crédito son determinados en el mercado de crédito. Este es un mercado donde la demanda proviene únicamente de las personas elegibles. Es un mercado no-walrasiano. El mercado de crédito no es como el mercado de papas. Por lo tanto, las ecuaciones (4.16) y (4.17) determinan la tasa de interés de equilibrio y la

cantidad de crédito intercambiado. Gráficamente, estas ecuaciones pueden ser representadas en la Figura 4.3, en la que el equilibrio de mercado tiene lugar en el punto G.

Un parámetro de la función de oferta es la tasa de interés internacional. Si esta tasa aumenta, la curva de oferta se desplazará hacia la izquierda. La tasa de interés de equilibrio también aumentará. Por lo tanto, hay un efecto positivo de la tasa de interés internacional sobre la tasa de interés doméstica. Un aumento en la tasa internacional cambiará el portafolio de los bancos y tendrá lugar una fuga de capitales. Como resultado, la oferta doméstica de crédito caerá. La tasa de interés doméstica aumentará. Este aumento tendrá un efecto secundario en los mercados de trabajo y dinero. El efecto final puede ser establecido resolviendo un sistema de equilibrio general. La tasa de interés internacional es una variable exógena en el mercado de crédito y, por lo tanto, en la totalidad de la economía.

No es necesario ir a la solución del nuevo sistema de equilibrio general. El efecto final de esta variable exógena en la solución de equilibrio general es fácil de seguir. Una mayor tasa de interés internacional causa un aumento en la tasa de interés doméstica, y eso desplaza la curva de demanda de trabajo hacia abajo. Luego, la tasa de interés doméstica podría ser sustituida por la tasa de interés internacional en la función de demanda de trabajo —ecuación (4.15) arriba—. Este efecto es similar al creado por una caída del coeficiente de términos de intercambio —caída que fue analizada anteriormente y presentada en la Figura 4.1—. El empleo, el producto y el salario real caen. Este modelo también predice que la tasa de salario real es procíclica.

En este modelo de equilibrio general con bancos, existen tres variables nominales que son la cantidad de dinero, el tipo de cambio y la tasa de interés. Pero existe un grado de libertad para elegir una de estas variables como exógena. Una vez hecha esta elección, las otras dos variables se determinarán endógenamente.

4.5. Resumen de la economía épsilon estática y abierta

El método de la estática comparativa ha producido un conjunto de relaciones de causalidad en las cuales los valores de las variables endógenas del sistema dependen de las variables exógenas. Las variables exógenas incluyen instrumentos de política doméstica (monetaria y fiscal), *shocks* externos (términos de intercambio y tasa de interés internacional), las dotaciones de recursos (*stock* de capital y cantidad de trabajo) y la desigualdad inicial en las dotaciones de recursos de los individuos.

Este conjunto de relaciones entre variables endógenas y exógenas del sistema, llamado proposiciones beta, necesita ser contrastado con los datos empíricos. El conjunto de relaciones del sistema de equilibrio general se llama el conjunto de *ecuaciones estructurales,* pero las proposiciones beta representan relaciones en la *forma reducida* del sistema, donde las variables endógenas dependerán solo de las variables exógenas.

Para las variables endógenas más importantes del sistema, tales como producto, empleo, nivel de ganancia y distribución, la forma reducida puede ser escrita como sigue:

$$
\begin{aligned}
Q_b^{o} &= F\,(K_b, S_h, S_m, z^*, r^*, \delta) \\
D_h^{o} &= G\,(K_b, S_h, S_m, z^*, r^*, \delta) \\
P^{o} &= H\,(K_b, S_h, S_m, z^*, r^*) \\
[P/W]^{o} &= J\,(\delta)
\end{aligned}
\qquad (4.18)
$$

Como en toda relación que muestra la forma reducida, las variables endógenas dependen solo de las variables exógenas del sistema. Así, los valores de equilibrio estático tanto del producto como del nivel de empleo dependen de todas las variables exógenas del sistema. En el nivel de ganancia no aparece la desigualdad inicial, pues si el tamaño del grupo de capitalistas se redujera, la ganancia total no cambiaría, excepto que menos gente la recibiría. En el caso de la distribución del ingreso nacional, el valor de equilibrio depende únicamente de la

desigualdad inicial en la dotación de activos económicos. La razón es que las demás variables exógenas no tienen un efecto definido o no tienen efecto alguno.

Los efectos de las variables exógenas en cada variable endógena se muestran en la matriz de proposiciones beta en el Cuadro 4.1. El efecto de cambios en la tasa de interés internacional r^* no se presenta, pero es equivalente al efecto inverso de un cambio en los términos de intercambio z, como se discutió en el texto. El efecto de un cambio en las dotaciones iniciales δ no se discute aquí, pero constituirá la cuestión central en los capítulos siguientes. Como se puede ver en este cuadro, ahora las predicciones empíricas de la teoría épsilon son consistentes con los hechos más destacados de los países del primer mundo, incluyendo el hecho de que el salario real y el empleo son todos procíclicos.

El producto neto que se obtiene en el proceso económico se llamará *ingreso nacional* (Y). Su nivel y su distribución pueden ser representados como sigue:

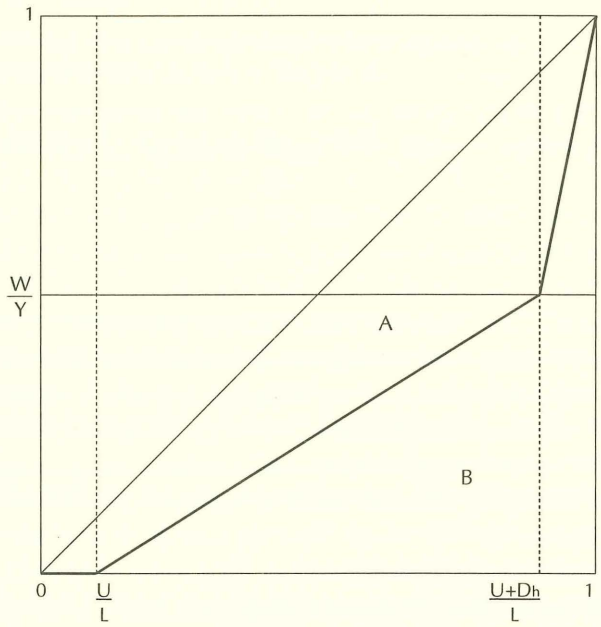
$$\begin{aligned} Y &= P + W \\ &= \kappa K + w D_h \\ S_h &= D_h + U \end{aligned} \qquad (4.19)$$

El ingreso nacional se distribuye entre capitalistas y trabajadores. Los beneficios pueden descomponerse en la tasa implícita de retorno del capital κ multiplicada por el *stock* de capital. Nótese que la tasa de retorno κ no es el precio de mercado de ningún bien o servicio; es la tasa de beneficio. Para dejar claro que el sistema capitalista opera con desempleo, la ecuación de asignación del trabajo incluye el tamaño del desempleo (U) con ingreso cero.

La concentración del ingreso nacional que resulta del proceso económico en épsilon se puede representar por una curva de Lorenz. Esta curva se presenta en la Figura 4.4. El eje horizontal mide la proporción de la población (L) que es desempleada, asa-

lariada y la que forma parte de la clase capitalista. Esta proporción se mide en forma acumulada y, por ello, la dimensión del eje es igual a 1. El eje vertical mide la participación de cada uno de esos grupos en el ingreso nacional (Y). La medición es también acumulada y la dimensión del eje debe ser igual a 1. La curva de Lorenz es la que aparece dibujada en segmentos gruesos. Esta curva muestra que la clase capitalista es relativamente pequeña, pero recibe una proporción relativamente alta del ingreso nacional. El conocido coeficiente de Gini reduce esta curva a un número, pues mide la proporción del área A sobre el triángulo que forman las áreas (A+B). Una curva de Lorenz más hacia fuera implica una mayor concentración del ingreso, y el correspondiente coeficiente de Gini será también mayor.

Figura 4.4
Curva de Lorenz sobre la distribución de ingresos en la economía épsilon



El concepto de ingreso nacional utilizado para la economía épsilon es equivalente a aquellos utilizados en la teoría neoclásica y keynesiana. El producto neto en la teoría clásica se llama *excedente económico,* e incluye beneficios y aquella parte de los salarios que está por encima del ingreso de subsistencia de los trabajadores, no el total de la masa salarial. Mediciones estadísticas del producto pueden variar según el punto de vista teórico; implícitamente, siempre hay una posición teórica incorporada en estas mediciones. Esta observación muestra que en el proceso de la falsación empírica se requiere de definiciones consistentes entre la teoría y la medición empírica.

Los resultados del modelo estático de la economía épsilon desarrollados en este capítulo y en el anterior pueden resumirse como sigue. El hecho estilizado es que —en el corto plazo— el salario real, el empleo y el producto real se mueven de forma conjunta en los países del primer mundo (Barro 1997). Esta relación empírica puede reproducirse con modelos de la teoría épsilon donde la curva de demanda de trabajo tenga parámetros adicionales al *stock* de capital.

Construyendo un modelo de economía abierta para la economía épsilon, dos variables exógenas han sido agregadas a la economía: el coeficiente de los términos de intercambio internacional y la tasa de interés internacional. Estas variables exógenas también constituyen parámetros de la curva de demanda de trabajo. Un aumento en los términos de intercambio causa un desplazamiento hacia arriba en la curva de demanda de trabajo. El mismo desplazamiento es originado por una caída en la tasa de interés internacional. La economía ve su producto, empleo y salario real incrementados. Una caída en los términos de intercambio o un aumento en la tasa de interés internacional tiene efectos en sentido contrario. Por lo tanto, los *shocks* internacionales (términos de intercambio y tasas de interés) causan un cambio en el equilibrio estático en la economía épsilon, donde este modelo abierto predice que el salario real y el nivel de empleo son todos procíclicos.

Dado que en la economía épsilon existe un grado de libertad para establecer como variable exógena sea la cantidad de dinero, el tipo de cambio o la tasa de interés, el gobierno puede elegir una de ellas como instrumento de política (y las restantes serán determinadas endógenamente). El cambio exógeno en cualquiera de estas variables nominales causa variaciones de corto plazo en las variables reales; en particular, el modelo predice cambios en el producto y empleo en las mismas direcciones (recesiones y expansiones), pero el salario real no es procíclico. La observación empírica de que el salario real es procíclico en el primer mundo indicaría el efecto neto de todos estos cambios en las variables exógenas.

Este modelo de la teoría épsilon ha utilizado el supuesto de que la inversión es exógena, pues se trata de un modelo estático. Si el *stock* de capital aumenta, la curva de demanda de trabajo se desplaza hacia fuera y genera un aumento en el nivel de empleo. Pero el desempleo no puede ser eliminado con la acumulación del capital. En la sociedad épsilon, el desempleo es una necesidad para el funcionamiento de la economía. Las predicciones de este modelo son consistentes con los datos básicos de los países del primer mundo. Por lo tanto, se puede concluir que los países del primer mundo se parecen a la economía épsilon, funcionan *como si* fuesen sociedades épsilon. No hay razón alguna para rechazar la teoría épsilon.

Capítulo 5
La sociedad omega

En los países del tercer mundo se observa que, entre los trabajadores que están fuera del mercado laboral, algunos están ocupados en la búsqueda de empleo en el mercado laboral, mientras que otros han logrado crear su propio empleo. Los que están autoempleados producen bienes y servicios en pequeñas unidades de producción que están localizadas en áreas rurales (en pequeñas fincas) y en áreas urbanas (en talleres pequeños, tiendas pequeñas, algunas de ellas ambulatorias). Estas pequeñas unidades coexisten con las grandes firmas domésticas y con las corporaciones multinacionales. Por el contrario, en el primer mundo, los trabajadores que están fuera del mercado laboral se encuentran en su gran mayoría buscando trabajo asalariado.

¿Cómo se puede explicar el funcionamiento de los países del tercer mundo? Construyendo una teoría, por supuesto. En este capítulo, se construirá una economía abstracta que busque capturar los elementos esenciales de estos países. Se supondrá que esta economía abstracta es sobrepoblada, y esta será su única diferencia respecto a la economía épsilon. Esta nueva sociedad abstracta será denominada la economía omega.

5.1. Una sociedad sobrepoblada

El concepto de sobrepoblación ha sido empleado en distintas formas en la literatura. La más común es aquella que considera que una economía está sobrepoblada si, y solo si, la productividad marginal de toda la oferta laboral es igual a cero o menor que el ingreso de subsistencia del trabajador (Geor-

gescu-Roegen 1967). Esta definición no es muy útil cuando se acepta que el mercado laboral es no-walrasiano y la economía capitalista debe operar con desempleo. ¿Es este exceso de oferta de trabajo una indicación de que la productividad marginal de la oferta laboral es igual a cero?

El problema analítico consiste en cómo distinguir entre el exceso de oferta de trabajo y la sobrepoblación. Se trata de encontrar una condición bajo la cual el exceso de oferta de trabajo fuese tan grande que la economía no podría operar como una economía épsilon. Esta condición puede ser la siguiente: las dotaciones de capital y trabajo en la economía son tales que cualquiera que sea la solución de mercado generará un exceso de oferta de trabajo considerable. ¿Cuán grande? El exceso de oferta de trabajo debe estar muy por encima del nivel de desempleo necesario para la disciplina laboral. El exceso de trabajo debe ser tan grande que los desempleados no podrían subsistir si se dedicasen solo a buscar trabajo. Con esa dotación de factores, esta economía no sería socialmente viable como una economía épsilon. En este estudio se adoptará esta definición de sobrepoblación.

Supóngase que, en una sociedad épsilon, la población aumentara exógenamente. Tal como se mostró en el capítulo anterior, el único efecto que tiene el aumento de la población sobre la economía es el aumento del desempleo. Si la población sigue aumentando, el desempleo se hará cada vez mayor, hasta que llegará un momento donde la economía ya no pueda operar bajo las reglas institucionales establecidas. Se produciría, entonces, un cambio cualitativo en la sociedad épsilon. Imagine que el Gobierno otorgara un subsidio de desempleo. Con este aumento de población y desempleo no sería viable, con el presupuesto público dado, seguir otorgando tal subsidio. En suma, la economía épsilon supone, para su funcionamiento, una relación capital-trabajo suficientemente elevada, pues de lo contrario no sería socialmente viable.

Esta definición de sobrepoblación es independiente del valor que tenga la productividad marginal de la oferta laboral. La so-

ciedad épsilon podría funcionar aun si este valor fuese cero; también podría funcionar si este valor fuese positivo. Las definiciones convencionales de sobrepoblación basadas en la productividad marginal del trabajo se han inscrito dentro de la teoría neoclásica. Claro, si el mercado laboral es walrasiano, el salario real se determinará donde se crucen las curvas de demanda y oferta; y cuanto mayor sea la oferta, el salario real walrasiano será más bajo. Así se puede llegar a una situación donde el salario real es demasiado bajo para cubrir las necesidades de subsistencia de los trabajadores. De aquí fue natural que se llegara a establecer la proposición de que «el sistema capitalista no puede operar bajo sobrepoblación». Si se acepta que el mercado laboral es no-walrasiano, la definición de sobrepoblación basada en la productividad marginal del trabajo deber ser abandonada. Eso es lo que se hace en este estudio

La economía omega es una economía sobrepoblada. Luego, en esta economía, el desempleo no puede ser usado como un mecanismo para extraer esfuerzo de los trabajadores. Como desempleados, los trabajadores se vuelven superfluos al funcionamiento del mercado laboral, y la rotación de empleos no podría generarles ingresos. Con el fin de hacer el capitalismo socialmente viable, el trabajo superfluo deberá producir sus ingresos fuera del sector capitalista. El supuesto es que los trabajadores generan sus propias unidades de producción. Estas unidades serán denominadas unidades de subsistencia, y su agregación se denominará *sector de subsistencia.*

Las proposiciones alfa para la economía omega incluyen las proposiciones que fueron establecidas para la economía épsilon y un nuevo supuesto: existe sobrepoblación en el mercado de trabajo. Para tener una referencia rápida, el conjunto de proposiciones alfa se expresa así:

$\alpha_{\omega}(1)$. *Contexto institucional.* Reglas: los individuos participan en los procesos económicos dotados de activos económicos; dichos activos están sujetos a derechos de pro-

piedad. Los individuos pueden intercambiar bienes sujeto a las normas sociales del mercado, que incluyen la norma de que los salarios nominales no pueden caer. Organizaciones: familias, firmas y el Gobierno.

$\alpha_{\omega}(2)$. *Condiciones iniciales.* Los individuos están dotados con cantidades desiguales de activos económicos. Existen dos clases sociales: los capitalistas y los trabajadores. Las dotaciones agregadas de capital y trabajo son tales que en la economía existe sobrepoblación.

$\alpha_{\omega}(3)$. *Racionalidad económica de los agentes.* Las motivaciones individuales son las del *homo economicus.* Los capitalistas tienen dos objetivos particulares: mantener su posición social y maximizar sus beneficios. En el mercado de trabajo, los trabajadores buscan maximizar su salario y minimizar su esfuerzo. Los capitalistas emplean mecanismos para extraer esfuerzo de los trabajadores.

$\alpha_{\omega}(4)$. *Relaciones de mercado.* Hay mercados walrasianos y no-walrasianos. El mercado de trabajo es no-walrasiano. No existe el mercado de alquileres servicios de bienes de capital. Los capitalistas no alquilan sus bienes de capital porque ellos prefieren las ganancias a las rentas.

5.2. Un modelo de la economía omega

En la economía omega solo se produce un bien (el bien B). La tecnología es tal que el bien puede ser producido con diferentes combinaciones de capital y trabajo. El producto también puede ser producido solamente con trabajo. El trabajo es esencial para la producción, pero el capital no lo es. Este es un supuesto distinto al de los modelos neoclásicos, donde tanto el capital como el trabajo son esenciales para la producción, esto es, donde el producto no puede ser producido únicamente con trabajo.

Hay dos tipos de unidades de producción, las firmas capitalistas y las unidades de subsistencia. La tecnología de producción

en el sector capitalista usa capital y trabajo, mientras que en el sector de subsistencia, solo emplea trabajo.[1] La productividad de trabajo en el sector de subsistencia es muy baja como para generar ahorros y acumulación de capital, ya sea en forma de capital humano o capital físico. De allí su denominación de sector de subsistencia. En un mundo de un bien, el producto en el sector de subsistencia se destina al auto-consumo, no al mercado.[2]

¿Cómo obtienen las firmas capitalistas la intensidad de trabajo deseada en una economía sobrepoblada? La tasa de desempleo no puede ser usada como mecanismo de disciplina, debido a que el exceso de oferta de trabajo es muy grande. Las firmas deben usar otro mecanismo.

Lewis (1954) desarrolló un modelo teórico para una economía sobrepoblada. Este trabajo clásico en la literatura se puede considerar como un modelo de la economía omega, como un primer modelo. Él introduce el supuesto de que el salario relativo es el mecanismo que emplean las firmas para obtener la intensidad de trabajo óptima. Los capitalistas pagan una tasa de salario que está por encima del costo de oportunidad de los trabajadores, que es igual al ingreso que pueden obtener en el sector de subsistencia. Este supuesto se puede escribir del siguiente modo:

$$w = (1 + c)\, v', \; c > 0 \qquad (5.1)$$

Aquí v' es la productividad marginal del trabajo en el sector de subsistencia; c es el coeficiente que mide el premio que debe

[1] Las unidades de subsistencia pueden tener *stock* de capital, pero estas unidades funcionan como si no lo tuviesen, porque la cantidad de capital es muy pequeña en relación con lo que se encuentra en el sector de los capitalistas. El sector de subsistencia es un subsistema de producción y consumo, no de acumulación de capital.

[2] Si el sector de subsistencia produce muchos bienes, este supuesto es equivalente a decir que la demanda de bienes producidos en el sector de subsistencia proviene del mismo sector; es decir, es como si el sector produjera para el auto-consumo.

recibir el trabajador para que respete la disciplina laboral. Este coeficiente depende del grado de conflicto social existente entre trabajadores y capitalistas, que está dado exógenamente. Con este salario, las firmas podrán extraer el nivel de esfuerzo óptimo a los trabajadores. Por lo tanto, este salario puede ser denominado el «salario de eficiencia». Lewis asumió, también, que la productividad media del trabajo en la economía de subsistencia es constante y, por lo tanto, igual a la productividad marginal del trabajo en el sector.

El modelo de Lewis se presenta en la Figura 5.1 (c). El segmento OO' mide la oferta de trabajo total. El empleo en el sector capitalista es medido desde el origen O y el autoempleo desde el origen O'. La productividad constante (marginal y media) del trabajo en el sector de subsistencia es igual a la línea horizontal empezando desde el punto v'. Dado el coeficiente c y la correspondiente intensidad del trabajo ($\lambda^*=1$), la productividad marginal del trabajo en el sector capitalista estará representada por la curva MR. En una economía competitiva, esta curva representa también a la curva de demanda de trabajo. Sea el salario resultante igual a w^o: a esta tasa salarial, el nivel de empleo en el sector capitalista es igual a OA, y el autoempleo en el sector de subsistencia es igual a AO'.

Si el salario real en el mercado de trabajo fuese igual al costo de oportunidad (esto es $w^o = v'$), el mercado de trabajo sería un mercado walrasiano. Pero, bajo esta forma, el mercado laboral no podría funcionar. El esfuerzo de los trabajadores sería inferior al que requieren las firmas, y la curva de productividad marginal se desplazaría hacia abajo y las ganancias serían menores.

En suma, en el modelo de Lewis, el nivel de productividad del trabajo del sector de subsistencia determina la tasa salarial. Dado el salario real, se establece el empleo en el mercado de trabajo; a esta tasa salarial, hay un exceso de oferta de trabajo. Todo el exceso de oferta de trabajo se desplaza hacia el sector de subsistencia.

Las predicciones que se obtienen del modelo de Lewis son inconsistentes con algunos datos básicos de la realidad del tercer mundo. En primer lugar, el modelo predice un nivel de desempleo igual a cero. Esta predicción es inconsistente con la realidad que enfrentan los países menos desarrollados, donde se observa desempleo, además de una cantidad significativa de personas autoempleadas. En segundo lugar, la ratio del salario real y el ingreso medio en el sector de subsistencia es fija. Este supuesto no parece ser consistente con los datos del mundo real. En tercer lugar, el salario real y el empleo están determinados ambos en el mercado de trabajo por variables únicamente reales como el *stock* de capital, la tecnología y la productividad del trabajo en el sector de subsistencia. Las variables nominales, tales como la oferta de dinero o el tipo de cambio, no juegan ningún rol. Si este modelo se incorporara a un sistema de equilibrio general, el modelo predeciría neutralidad del dinero, pero empíricamente el dinero es no neutral.

Se requiere, entonces, un nuevo modelo para la economía omega. Se introducen ahora dos supuestos auxiliares. En primer lugar, se supone rendimientos decrecientes en el sector de subsistencia. Las unidades de producción en este sector tienen distintas productividades debido no a diferencias en el conocimiento tecnológico, sino a diferencias en la ubicación espacial de las unidades, que crean diferencias en el costo del acceso a bienes públicos y recursos naturales. Los rendimientos decrecientes se producen cuando entran más unidades al sector de subsistencia, pues los nuevos se ubicarán en espacios de menor calidad económica. Hay que notar que este concepto de rendimientos decrecientes es el concepto ricardiano, no el de la teoría neoclásica.

En el sector de subsistencia, la curva de productividad marginal representa la curva de oferta de trabajo. A un salario dado en el mercado laboral, los trabajadores tienen la opción de elegir entre desempleo o autoempleo. Los trabajadores que puedan obtener un ingreso más elevado siendo autoempleados optarán por no entrar en el mercado de trabajo. Si el salario se

eleva, disminuirá la cantidad de trabajadores que puedan obtener ingresos superiores como autoempleados (debido a los rendimientos decrecientes) y más trabajadores optarán por ingresar al mercado laboral. La oferta de trabajo está dada por la productividad marginal del trabajo y tiene pendiente positiva cuando es medida desde el origen O. Esto se aprecia en la curva rm en la Figura 5.1 (c).

Si el salario real se determina donde las curvas de oferta y demanda se cruzan, la situación de equilibrio se producirá en el punto G. Este equilibrio implica que el mercado de trabajo es walrasiano. Pero el mercado de trabajo no podría operar de esta manera debido a que disminuiría la intensidad de trabajo y, entonces, la curva de productividad marginal de trabajo se desplazaría hacia abajo y las ganancias disminuirían.

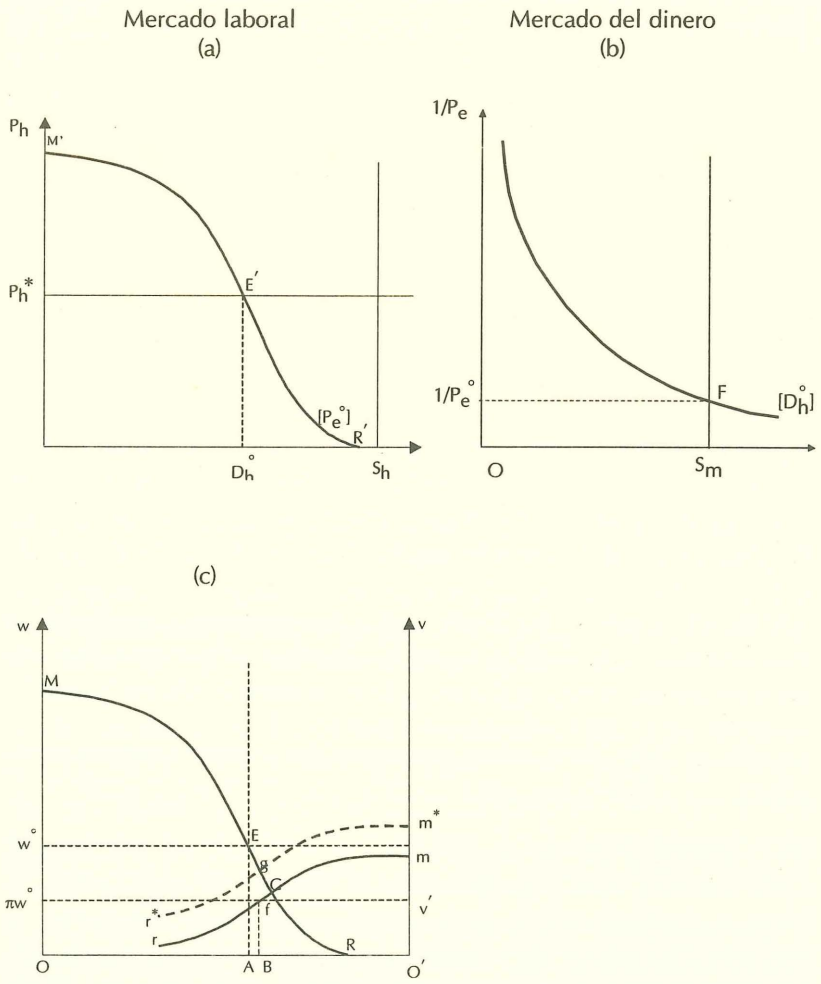
Se introducirá ahora el supuesto de Lewis respecto a que la intensidad del trabajo depende positivamente de la diferencia entre el salario real y el costo de oportunidad del trabajo. La curva de extracción del esfuerzo salarial es obtenida al desplazar hacia arriba la curva de oferta laboral en una proporción (1+ c). Esto se muestra como la curva m^*r^* en la Figura 5.1 (c). En tanto esta brecha prevalezca, el grado de esfuerzo laboral será λ^* y la curva de demanda de trabajo permanecerá en la misma posición. Esta brecha puede operar como una restricción para la determinación del salario. Esta curva es, en efecto, la curva de extracción de esfuerzo.

El segundo supuesto auxiliar es que, aunque la racionalidad de los capitalistas no difiere de la que se supuso en la economía épsilon, ahora sus restricciones son distintas. Las firmas capitalistas en la economía omega —en un contexto de sobrepoblación— tienen la siguiente racionalidad:

$$\text{Maximizan} \quad P' = P_b Q_b - P_c D_c - P_h D_h, \tag{5.2}$$

$$\text{sujeto a} \quad Q_b = F(D_h, K_b),\ \lambda^* = 1$$
$$= D_c / n$$
$$w \geq w^* = (1 + c)\, v'$$

Figura 5.1
El equilibrio general en la economía omega



Entre las restricciones para la conducta de los capitalistas se incluye ahora la condición de que el salario en el mercado de trabajo debe ser mayor al costo de oportunidad de los trabajadores. Si los trabajadores no respetan la disciplina laboral, ellos sufrirán un costo económico. Dada la curva de extracción del esfuerzo y la curva de demanda de trabajo, hay un límite inferior para el salario real (w^*), determinado donde las dos curvas se cruzan (el punto g de la Figura 5.1). Este salario de eficiencia implica a su vez el nivel máximo de empleo (L^*).

La lógica individual de los trabajadores puede ser resumida como sigue. El trabajador opera en un contexto donde existe exceso de oferta de trabajo pero busca empleo en el sector capitalista, debido a que el salario es mayor que el ingreso que pudiese obtener en el sector de subsistencia. El sistema de incentivos lo conduce a este comportamiento, pero no hay trabajo para todos, no importa cuánto un trabajador se esfuerce en la búsqueda de empleo. Debido a que existe un porcentaje de vacantes como trabajadores en el sector capitalista por los despidos de aquellos encontrados en falta disciplinaria, entonces hay una probabilidad de encontrar trabajo en este sector. Supóngase que la probabilidad es la misma para todos los trabajadores.

Sea w^e el salario esperado de los trabajadores cuando están buscando un trabajo. Para una probabilidad exógenamente dada para encontrar empleo (π), el salario esperado depende positivamente de la tasa salarial del mercado. Este supuesto implica un salario uniforme esperado para todos los trabajadores. Este puede ser escrito como:

$$w^e = \pi w, \text{ donde } 0 < \pi < 1 \qquad (5.3)$$

Los trabajadores excluidos del mercado laboral evaluarán el salario esperado cuando están en una activa búsqueda de trabajo (desempleados) frente al ingreso seguro que puedan obtener en el sector de subsistencia. Si el salario real esperado es mayor que el ingreso seguro en el sector de subsistencia, los trabajadores optarán por buscan un ingreso salarial y continua-

rán desempleados; si el salario real esperado es menor, ellos optarán por ser autoempleados en el sector de subsistencia. Debido a que el salario esperado depende del salario de mercado, cuanto mayor sea el salario de mercado, mayor es la proporción de desempleados; y, dado un exceso de oferta laboral (E_h), menor será la proporción de autoempleados.

Los trabajadores que son excluidos del empleo asalariado tienen así la opción de elegir entre desempleo y autoempleo. Pero es una elección luego de ser excluidos, es decir, es una elección en una situación de segunda opción.

5.3. Equilibrio general

Considere una economía omega abierta y estática. Existen cuatro mercados que operan bajo competencia perfecta. Estos son el mercado del bien B, el mercado de trabajo, el mercado del dinero y el mercado de divisas. Las condiciones de equilibrio de estos mercados son las siguientes:

$$\begin{array}{lll} \text{Mercado de trabajo} & : & S_h = D_h + E_h \\ \text{Mercado de bienes} & : & S_b = D_b \\ \text{Mercado de divisas} & : & P_b^* X_b = P_c^* D_c \\ \text{Mercado de dinero} & : & S_m = D_m, \end{array} \qquad (5.4)$$

sujeto a la restricción del presupuesto agregado:

$$P_b (D_b - S_b) + P_e (P_c^* D_c - P_b^* X_b) + (D_m - S_m) = 0$$

El mercado de trabajo es no-walrasiano, donde E_h indica el exceso de oferta de trabajo. El resto son mercados walrasianos.

Las condiciones de equilibrio general son similares a aquellas que se establecieron para la economía épsilon. El mercado de divisas está siempre en equilibrio, pues este intercambio está ya incorporado en las restricciones presupuestarias de las firmas. Esto nos deja con dos mercados walrasianos en el que uno puede ser eliminado. Luego, el equilibrio general puede ser esta-

blecido resolviendo dos mercados, el mercado de dinero y el mercado de trabajo. Estas ecuaciones son:

$$D_h = H(P_e; P_h, P_b^*, P_c^*, K_b),\ H_1>0,\ H_2<0,\ H_3>0,\ H_4<0,\ H_5>0$$
$$S_m = M(P_e, D_h; P_h, P_b^*),\ M_i>0, \text{ para todo i,} \qquad (5.5)$$
$$\text{sujeto a } w \geq w^* = (1+c)\,v'$$

Este sistema es idéntico al establecido en la economía épsilon. Hay dos ecuaciones que pueden resolver las dos variables endógenas: el tipo de cambio P_e y el nivel de empleo D_h.

La Figura 5.1 muestra la solución de equilibrio general. En la Figura 5.1 (a), se muestra la curva de demanda de trabajo en términos del salario nominal. La posición de esta curva queda determinada una vez que el tipo de cambio es conocido. En la Figura 5.1 (b), se muestra el mercado de dinero, donde la posición de la curva de demanda es determinada una vez que se conoce el nivel de empleo. Así, hay un par de valores que determinan el valor de equilibrio de estas variables endógenas. En la Figura 5.1 (c), se muestra el mercado de trabajo en términos de los salarios reales. La demanda de trabajo está dada por la curva MR. El nivel de empleo de equilibrio es OA. Esta cantidad de empleo corresponde al salario real de equilibrio w^o. El exceso de la oferta de trabajo es igual a AO', donde OO' es el total de oferta de trabajo.

La solución es secuencial. Primero, se determina el salario real y el empleo en el sector capitalista. Una vez que estos valores son conocidos (w^o, D_h^o), la asignación del exceso de oferta de trabajo (E_h^o) entre desempleo (U) y autoempleo (L_s) es determinada. Las condiciones para esta solución incluyen:

$$E_h^o = L_s + U$$
$$V = J(L_s),\ J'>0,\ J''<0 \qquad (5.6)$$
$$\pi w^o = J'(L_s)$$

La primera ecuación muestra que el exceso de oferta de trabajo en la economía omega incluye el desempleo y el autoempleo.

Esto es distinto en la economía épsilon, donde el exceso de oferta de trabajo toma únicamente la forma de desempleo. La segunda ecuación es la función de producción en el sector de subsistencia, sujeto a rendimientos decrecientes.

La condición de equilibrio para la asignación de trabajo entre autoempleo y desempleo está establecida en la tercera ecuación. Debido a que el costo de oportunidad de buscar un trabajo es la productividad marginal del trabajo en el sector de subsistencia, la función $J'(L_s)$, ambos ingresos deben ser iguales en equilibrio. Dado el salario esperado, los trabajadores que pueden obtener un ingreso más alto en el sector de subsistencia optarán por el autoempleo. Los trabajadores que podrían obtener un menor ingreso en el sector de subsistencia optarán por buscar un empleo como asalariados y preferirán estar desempleados.

De este modo, la igualdad entre el salario esperado y el valor de la productividad marginal del trabajo en el sector de subsistencia es la línea divisoria que determina la asignación del trabajo entre el desempleo y el autoempleo. El equilibrio general de esta economía asegura que aquellos que reciben un salario tendrán un ingreso más elevado que aquellos que están excluidos del mercado de trabajo.[3]

---

[3] El periodo de trabajo en cada sector no tiene por qué ser el mismo en ambos sectores. Roemer (1982) emplea un modelo en el cual esta igualdad es una condición de equilibrio. Bajo condiciones de competencia en el mercado de trabajo y con una productividad de trabajo media constante, con un nivel mayor en el sector capitalista, el equilibrio se alcanza cuando los trabajadores trabajan el mismo número de horas como asalariados o como auto-empleados. Si las firmas fijan un número menor de horas, todos los trabajadores buscarán un empleo asalariado; y si fijan un número mayor de horas, ningún trabajador buscará empleo asalariado. Pero, entonces, todos los trabajadores obtienen el mismo ingreso, que es igual a la productividad media del sector de subsistencia. Por lo tanto, Roemer trata el mercado de trabajo como el mercado de papas: este mercado se limpia y no se requiere de ningún mecanismo de extracción del excedente de los trabajadores. En nuestro modelo, la condición de equilibrio es tal que el salario real debe ser mayor que el ingreso marginal en el sector de

Este equilibrio se puede ver en la Figura 5.1 (c). La oferta de trabajo total es igual a OO'. La curva de demanda de trabajo está dada por la curva MR. La productividad marginal del trabajo en el sector de subsistencia está representada por la curva *mr,* medido desde el origen O'. La curva m^*r^* representa la restricción del salario real, que dice que el salario real debe ser mayor que el costo de oportunidad del trabajo. El salario real de equilibrio (w^o) y la cantidad de trabajo demandada (OA) son determinados en el sector capitalista. El exceso de oferta de trabajo (AO') también es determinado.

Ahora que el salario es conocido, el salario esperado queda determinado. Está denominado como una fracción fija del salario de mercado. Este valor debe ser igual a la productividad marginal del trabajo, y eso se produce en el punto *f.* A este nivel de autoempleo, los trabajadores pueden hacer ingresos mayores que los del salario esperado. Más allá de este nivel, el salario esperado es mayor que el ingreso marginal. El exceso de oferta de trabajo (AO') se asigna al desempleo (AB) y al autoempleo (BO').

La brecha entre el salario real y la productividad marginal del total de trabajadores autoempleados debe satisfacer la restricción de que la diferencia debe ser igual o mayor que la brecha necesaria para asegurar la disciplina laboral. En el panel (c), el salario real de equilibrio está por encima de la curva m^*r^*. Si esta condición no es satisfecha, los trabajadores tendrán incentivos para incumplir con la disciplina laboral y las ganancias podrían caer; por lo tanto, las firmas buscarán incrementar el salario nominal con el fin de incrementar el salario real y mantener la intensidad de trabajo y sus ganancias. Si no se da esta brecha, el equilibrio general no será posible.

El mercado de trabajo es claramente un mercado no-walrasiano. Los trabajadores están deseosos de intercambiar su

subsistencia. La extensión de la jornada de trabajo por sectores puede ser diferente debido a la existencia generalizada de rendimientos decrecientes.

trabajo en el mercado real, pero no pueden hacerlo. El desempleo es el resultado de una elección económica hecha por aquellos trabajadores excluidos del mercado de trabajo. Este desempleo es «voluntario» en el sentido de que es el resultado de una elección económica (como a veces la mortalidad infantil es una triste elección económica hecha por las familias pobres). Pero es involuntario en el sentido de que esta no es la situación preferida por los trabajadores; el desempleo es una elección en una situación de segunda opción.

Como se puede ver en la Figura 5.1, el valor de equilibrio en el desempleo no necesariamente tiene que ser un número positivo, puede ser cero. En una economía con trabajo abundante, el desempleo no es una necesidad para el funcionamiento del sistema capitalista; en contraste, para la economía épsilon sí lo es.

El ingreso nacional (Y), su distribución y la asignación de la oferta de trabajo se pueden representar por las siguientes ecuaciones:

$$\begin{aligned} Y &= P + W + V \\ &= P + w D_h + v L_s \\ S_h &= D_h + L_s + U \end{aligned} \qquad (5.7)$$

Mientras que el salario real *w* es igual a la productividad marginal del trabajo en el sector capitalista, el ingreso medio de los trabajadores en el sector de subsistencia es igual a la productividad media del trabajo *v*. Tanto el nivel de ingreso nacional como su distribución son endógenas.

En la Figura 5.1, el ingreso nacional es igual a la suma del área debajo de la curva ME (el producto total del sector capitalista) más el área debajo de la curva *mf* (el producto total en el sector de subsistencia). Las firmas capitalistas y las unidades en el sector de subsistencia producen el mismo bien. En este mundo de un solo bien, esta agregación está plenamente justificada.

La solución de equilibrio general es claramente ineficiente. La asignación del trabajo dada por la intersección de las curvas

de oferta y de demanda es la que maximiza el ingreso nacional. Este equilibrio ocurre en el punto G en la Figura 5.1 (c), donde la curva de oferta de trabajo está dada por la curva *rm,* medida desde el origen O. Pero esta solución no es alcanzable económicamente, pues requeriría de un mercado de trabajo walrasiano. La economía omega es, por tanto, inherentemente ineficiente.

El equilibrio en la distribución de ingresos puede verse en la Figura 5.1 (c). El área debajo de la curva ME es el producto total generado en el sector capitalista, distribuido entre las ganancias (el área w^0ME) y los salarios (el área Ow^0EA). El área debajo de la curva *mf* es el producto total generado en el sector de subsistencia. En suma, la desigualdad es resultado de la presencia de dos sectores en el proceso económico. Existe desigualdad dentro del subsistema capitalista, entre los capitalistas y los asalariados; y también existe desigualdad entre los diferentes segmentos de los trabajadores: aquellos que reciben un salario homogéneo y aquellos que ganan ingresos menores en el sector de subsistencia, y aquellos que están desempleados y que no perciben ningún ingreso.

Dado que la mano de obra es homogénea, ¿podría persistir esta desigualdad? La pobreza está formada por los trabajadores que están excluidos del mercado laboral. Pero ellos podrían ofrecer su trabajo a un salario nominal inferior al del mercado para lograr el empleo asalariado. ¿Por qué no lo hacen? Las normas sociales impiden esta competencia entre trabajadores. Existe, además, un límite en el nivel del empleo asalariado dado por el salario de eficiencia. En consecuencia, los trabajadores excluidos del mercado laboral no tienen otra elección que o buscar empleo como asalariados o autoemplearse. Pero debe quedar en claro que esta es una elección bajo una situación de segunda opción.

Los trabajadores excluidos del mercado laboral también podrían crear firmas o expandir el tamaño de las unidades donde se encuentran autoempleados. Esta línea de acción implica la acumulación de capital —lo que, a su vez, requiere de financiamiento—. Pero nada de esto es viable porque, como se

mostró en la economía épsilon, estos trabajadores también están excluidos de los mercados de crédito y de seguros. Por lo tanto, la desigualdad mostrada en el equilibrio general es, en efecto, una situación de equilibrio: nadie tiene el poder y el incentivo para cambiar esta situación.

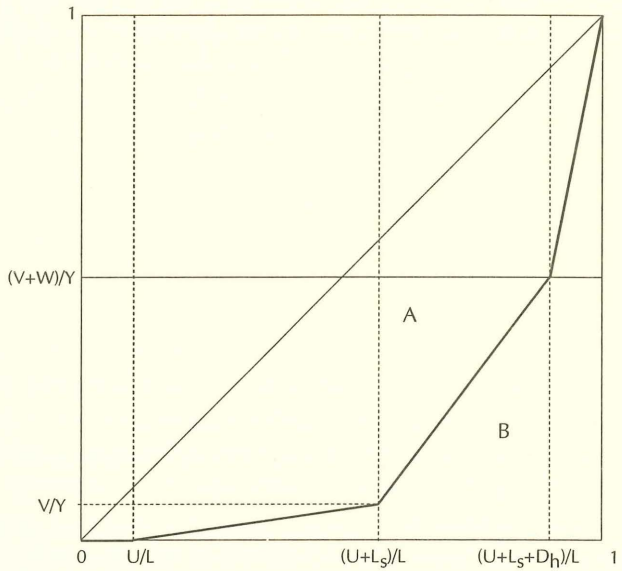
En la economía omega, en suma, el sistema de incentivos para obtener la intensidad de esfuerzo óptima, que implica obtener las máximas ganancias, está basado en el salario relativo (mientras que en la economía épsilon, está basado en el desempleo). El equilibrio general estático de la economía omega es determinado secuencialmente: una vez que el equilibrio en el subsistema capitalista queda determinado, el producto y el empleo en el sector de subsistencia se determinan juntamente con el desempleo. El empleo y el ingreso en el sector de subsistencia son resultado de la solución de mercado. En el sistema estático, estos valores de equilibrio general se repiten periodo tras periodo, si las variables exógenas permanecen constantes.

La concentración del ingreso que resulta del proceso económico se puede representar por una curva de Lorenz. En la Figura 5.2 se muestra la proporción de cada grupo de trabajadores y de la clase capitalista en la población y su respectiva participación en el ingreso nacional. La clase capitalista tiene una participación en el ingreso nacional que es mucho mayor comparada a su tamaño relativo.

5.4. Proposiciones beta

El equilibrio general estático que se ha establecido en la economía omega es secuencial. El equilibrio en el sector capitalista se determina en primer lugar y puede, en consecuencia, ser analizado por separado. En realidad, la condición de equilibrio general en el modelo omega, ecuación (5.5), es la misma condición presentada en el modelo épsilon del capítulo IV —la ecuación (4.9)—. Las restricciones parecen distintas, pero se las puede

Figura 5.2
Curva de Lorenz sobre la distribución de ingresos en la economía omega



expresar en términos de nivel de empleo máximo L^*. Por lo tanto, el sistema de relaciones entre la curva HH y la curva MM presentado en el capítulo IV, en la Figura 4.2, puede ser aplicado al modelo omega presentado aquí. La única diferencia es que aquí L^* se refiere al empleo máximo que es consistente con el salario de eficiencia w^* (y no con la tasa de desempleo de eficiencia u^*).

El equilibrio es único y estable. La estática comparativa puede entonces ser aplicada a este equilibrio para obtener proposiciones beta. Debido a que la solución de equilibrio general es secuencial, las relaciones de causalidad operan para cada subsistema.

En el sector capitalista, las variables son clasificadas como sigue:

Variables endógenas: P_e, D_h, P_h, P_b, Q_b, w, W, P, E_h
Variables exógenas : S_m, z^*, r^*, S_h, K_b, δ

Para el sector capitalista las proposiciones beta son derivadas del sistema de mercado. La matriz de las correspondientes proposiciones beta está presentada en el Cuadro 5.1.

El modelo de omega que se ha presentado en la Figura 5.1 incluye el sector capitalista, que opera como en la economía épsilon, para el caso donde $u>u^*$. Las proposiciones beta del modelo estático y abierto de la economía épsilon han sido desarrolladas ya en el capítulo 4 y están en el Cuadro 4.1. Estos resultados también se aplican al modelo de omega y están representados en la parte superior del Cuadro 5.1.

En el sector de subsistencia, las variables son clasificadas como sigue:

Variables endógenas: L_s, U, v, V
Variables exógenas : w, D_h, E_h

Las variables exógenas del sistema completo afectan a las variables endógenas del sector de subsistencia vía su efecto en la solución del mercado de trabajo. El salario real y el exceso de oferta de trabajo son variables endógenas en el sistema capitalista pero son exógenas en el sector de subsistencia, así como en la determinación del nivel de desempleo.

La primera situación ocurre cuando la curva de demanda de trabajo permanece fija. El modelo de épsilon mostró que un incremento en la oferta de dinero causa una caída en el salario real e incrementa el nivel de empleo. Entonces se reduce el exceso de oferta de trabajo. Por lo tanto, hay que considerar dos efectos. El efecto del salario real causa un alza en el nivel de autoempleo. La caída en el salario real hace que sea menos rentable para el trabajador permanecer desempleado; y, dado el exceso de oferta de trabajo, el desempleo cae. La disminución del exceso de oferta de trabajo tiene el efecto de reducir el

Cuadro 5.1
Matriz de proposiciones beta para un modelo omega estático ($D_h < L^*$)

Variables endógenas	Variables exógenas			
	S_m	z^*	K_b	S_h
P_e	+	−	−	0
D_h	+	+	+	0
Q_b	+	+	+	0
w	−	+	+	0
W	?	+	+	0
P	+	+	?	0
P_b	+	−	−	0
P_h	0	0	0	0
L_s	+	−	−	0
v	−	+	+	0
V	+	−	−	0
U	−	?	?	+
Y	+	+	+	0
P/W	?	?	?	0
V/Y	?	−	−	0
w /v	?	?	?	0

desempleo debido a que el autoempleo es determinado por el salario real, que permanece constante. Entonces, el efecto combinado de una caída en el salario real y una caída en el exceso de oferta de trabajo ocasiona un incremento en el nivel de autoempleo y una caída en el nivel de desempleo.

A medida que aumente la oferta monetaria, el nivel de precios y el nivel de empleo seguirán aumentando. Pero, debido a que existe en el sistema un nivel de empleo máximo, el aumento en la cantidad de dinero hará que se llegue a ese tope de empleo. Una mayor oferta de dinero cuando el nivel de empleo haya llegado al nivel de empleo máximo tendrá el efecto de aumentar solo el nivel de precios. Las variables reales ya no se modifica-

rán. El dinero se vuelve neutral. Este modelo de omega predice neutralidad del dinero. El Cuadro 5.1 muestra las proposiciones beta para el caso donde el equilibrio nuevo en el nivel de empleo se da todavía en el rango debajo del empleo máximo.

La segunda situación a ser considerada se refiere al desplazamiento de corto plazo de la curva de demanda de trabajo. El modelo de épsilon abierto mostró que los *shocks* externos, tales como el cambio de los términos de intercambio o cambios en la tasa de interés internacional, generan ese desplazamiento. Una proposición beta dice que un incremento en los términos de intercambio causa un incremento en el salario real y en el empleo. El efecto del salario real reduce el autoempleo, y esto incrementa el número de trabajadores que pueden encontrar rentable la búsqueda de empleos, esto es, aumenta el desempleo. La caída de la oferta de trabajo reduce el desempleo. El efecto combinado es que el autoempleo cae, pero el cambio en el desempleo es indeterminado.

El aumento en la oferta laboral no tienen ningún efecto sobre el sector capitalista. El único efecto es el de aumentar el exceso de oferta laboral y, más precisamente, el de aumentar el nivel de desempleo.

Hasta ahora la curva de productividad marginal del sector de subsistencia ha permanecido constante. Si esta curva se elevara de manera exógena, la curva de extracción de esfuerzo también se elevaría. En la Figura 5.1, la curva m^*r^* se desplazaría hacia arriba. Se darían dos posibles resultados. Primero, considérese que el punto E todavía se encuentra por encima de la nueva curva. En este caso, el autoempleo sube y el desempleo baja, pero nada cambia en el sector capitalista. Segundo, considere que el punto E se encuentre ahora por debajo de la nueva curva. En este caso, el salario relativo estaría por debajo del necesario para mantener la disciplina laboral. El salario nominal tendría entonces que subir para restablecer el salario real de eficiencia. Un continuo aumento en la productividad del sector de subsistencia tendría, finalmente, el efecto de aumentar el salario real y reducir el nivel de desempleo.

Este resultado es diferente al que se obtiene con el modelo de Lewis. Allí cualquier aumento en la productividad del sector de subsistencia tiene el efecto de elevar el salario real. En el modelo presentado aquí, este efecto ocurre solamente para aumentos significativos en la productividad del sector de subsistencia.

La tercera situación a considerar se refiere al desplazamiento de largo plazo de la curva de demanda de trabajo. Este desplazamiento proviene de la inversión, de la acumulación de capital. Como resultado, el nivel de empleo y el salario real aumentan. Además, hay que notar que el valor que toman tanto el salario de eficiencia como el nivel de empleo máximo depende del *stock* de capital. Luego, ambos valores aumentan con la acumulación de capital.

Si la acumulación de capital es continua, la economía dejará de ser sobrepoblada. El exceso de oferta laboral se reducirá continuamente y el salario real aumentará también continuamente. El sector de subsistencia se reducirá. La economía omega sufrirá una transformación social y devendrá en una economía épsilon. La tasa de desempleo sustituirá al salario relativo como mecanismo de la disciplina laboral.

La distribución de ingresos en la economía omega se refiere a la participación de P, W y V en el ingreso nacional. El efecto de las variables exógenas sobre las ratios P/W, V/Y y w/v es indeterminado o nulo, tal como se puede ver en el Cuadro 5.1. Aunque estos resultados se pueden derivar con facilidad, merece discutirse el efecto de un desplazamiento en la curva de demanda de trabajo (sea por un aumento en el coeficiente de los términos de intercambio o por un aumento en el *stock* de capital). Como se mostró arriba, el efecto es un aumento en el empleo y en el salario real en el sector capitalista. La participación del sector de subsistencia en el ingreso nacional disminuye. Pero hay que notar que esta disminución se debe a una caída en el autoempleo y no a una caída en el ingreso medio del sector, que más bien aumenta. El efecto sobre la desigualdad agregada es ambiguo también en este caso. En suma, el grado de desigualdad inicial tiende a persistir.

5.5. Consistencia empírica: los países del tercer mundo

¿Se parecen los países del tercer mundo a la economía omega? Como lo muestran las publicaciones de estadísticas internacionales (la Organización Internacional del Trabajo y el Banco Mundial), la estructura del empleo y la estructura del ingreso nacional en el tercer mundo parece ser consistente con las proposiciones beta de la economía omega.

La economía omega debe mostrar no solo menores niveles de ingreso per cápita sino, también, menores proporciones de asalariados dentro de la fuerza laboral en comparación con la situación en los países del primer mundo. Para una muestra de ocho países del primer mundo y diez países de Latinoamérica, sobre la base de los datos de OIT, este autor calculó que el promedio de la ratio de asalariados en relación con el total de la fuerza es de 84% en el primer mundo y 59% en Latinoamérica en 1996.[4] La situación de Latinoamérica está posiblemente por encima del promedio de la del tercer mundo.

Las estadísticas internacionales también muestran que, en los países del tercer mundo, los trabajadores que no son asalariados se distribuyen entre desempleados y autoempleados. Los autoempleados constituyen significativas proporciones de la fuerza laboral y sus ingresos son parte importante del ingreso nacional. Este es el grupo en el que la incidencia de la pobreza es mayor.

La práctica común de emplear la tasa de desempleo como criterio para hacer comparaciones internacionales del exceso de oferta de trabajo es teóricamente inaceptable. El exceso de oferta de trabajo en el primer mundo está dado solamente por el desempleo. Pero en el tercer mundo, el desempleo representa solamente una parte del exceso de oferta laboral y, usualmente,

[4] La sub muestra incluía Japón, Dinamarca, Alemania, Estados Unidos, Francia, Suecia, España y Portugal como Primer Mundo, y México, Brasil, Panamá, Costa Rica, Perú, Colombia, El Salvador, República Dominicana, Ecuador, y Bolivia para Latinoamérica. La fuente fue la siguiente: OIT. *World Employment Report 1997*. Cuadros 2A, 2D, y 2E.

la parte más pequeña. En el tercer mundo, el nivel de desempleo subestima de manera significativa el exceso de oferta laboral.

Aunque los estudios estadísticos sobre las relaciones agregadas en la economía de los países del tercer mundo son escasos, se puede señalar sobre la base de una información menos sistemática que las siguientes predicciones de la teoría omega son consistentes con esta realidad. Estas predicciones son las siguientes:

a. En el largo plazo, la inversión, el salario real y el nivel de empleo asalariado están positivamente correlacionados.
b. En el corto plazo, tanto el salario real como el empleo asalariado muestran variaciones que son procíclicas.

Históricamente, la teoría omega es consistente con el proceso de desarrollo que han seguido los países del primer mundo. En general, estos países empezaron como economías omegas y, vía el proceso de acumulación de capital, se convirtieron en economías épsilon. Pero esta no es la historia de los países del tercer mundo. Muy pocos países han dejado de pertenecer al tercer mundo y se han unido al club del primer mundo. Como ha mostrado el historiador económico Maddison (1995), los miembros del grupo de países ricos de 1820 siguen siendo los miembros del club de países ricos de hoy, con la única excepción de Argentina —que dejó el club— y Japón —que es el único miembro nuevo del club y el único caso de éxito en 150 años—. Corea del Sur todavía no es miembro del club pero está en proceso de alcanzarlo.

No importa cómo se hagan los cálculos y se escojan los periodos para encontrar países que constituyan casos de éxito, o en camino a ser exitosos, el hecho es que estos casos constituyen la excepción y no la regla. La regla general es que no hay convergencia en el nivel de ingreso entre los países del primer mundo y los del tercero (Barro y Sala - i - Martin 1995). Esta evidencia empírica es inconsistente con las predicciones derivadas de la teoría omega. La mayoría de los países del tercer mundo no se parecen a la economía omega. Se necesita una nueva teoría. Esta es el tema del capítulo siguiente.

Capítulo 6
La sociedad sigma

Las teorías económicas convencionales suponen sociedades en las que todos los individuos son homogéneos en todo sentido, excepto en sus dotaciones iniciales de activos económicos. La desigualdad en la dotación de otros activos son simplemente ignorados. Las teorías épsilon y omega desarrolladas en este estudio también han hecho este mismo supuesto. El resultado ha sido que la teoría omega no puede explicar los rasgos fundamentales de la economía mundial. Ese supuesto será abandonado ahora. Una nueva sociedad abstracta —la sociedad sigma— en la que también existe desigualdad en la dotación inicial de activos sociales será construida y desarrollada en este capítulo.

6.1. Una sociedad heterogénea y jerárquica

Sigma es una sociedad en la que los individuos participan en el proceso económico dotados de activos económicos y sociales. Los activos económicos incluyen varias formas de capital: físico, humano y financiero. El capital físico y financiero está concentrado en un pequeño grupo social: la clase capitalista.

Los activos sociales incluyen activos políticos y activos culturales. Los activos políticos se definen como la capacidad de un individuo para ejercer derechos. De aquí se sigue que la ciudadanía es un activo político, que da lugar a derechos y obligaciones. Debido a la desigualdad en la distribución de activos políticos, en la sociedad se crea una jerarquía de ciudadanos; aparecen así ciudadanos de primera y segunda categoría. No todos son iguales frente a la ley y, por lo tanto, los grupos ubica-

dos en los niveles más bajos de la jerarquía tienen un acceso relativamente limitado a los derechos económicos establecidos por la sociedad. Los derechos económicos toman la forma de bienes públicos, tales como servicios de educación y salud, y sistemas de protección social.

Sigma es una sociedad multiétnica. Los individuos pertenecen a diversos grupos étnicos; dichos grupos se diferencian entre sí por un origen distinto en cuanto a raza, nacionalidad, religión, lengua o costumbres. Estos elementos étnicos no tienen la misma valoración en la sociedad. Están ordenados en una jerarquía social, históricamente construida. Esa valoración otorga a los individuos activos sociales, según el grupo de afiliación, que en este caso denominaremos activos culturales. Se les llamará así porque se basan en creencias que son aprendidas y transmitidas de generación en generación.

Los activos culturales proporcionan a las personas ya sea prestigio o estigma social —que conduce a la discriminación y segregación—. Esta valoración desigual de los activos culturales implica la existencia de grupos con distinta posición social en la sociedad. Las personas no son como átomos, pues están socialmente situados; así, pueden nacer en la familia equivocada o en el lugar equivocado. En resumen, mientras que los activos económicos indican lo que una persona tiene, los activos políticos y culturales indican lo que una persona es.

En sigma, la desigualdad en activos económicos da lugar a la existencia de clases sociales y la desigualdad en activos sociales a la existencia de infraclases (*the underclass*) dentro de la estructura social. Las infraclases lo conforman grupos sociales con escasa dotación de activos sociales. Sigma es una sociedad de clases y de etnias; es una sociedad más compleja que epsilon y omega.

Al revés de los activos económicos, los activos políticos y culturales son intangibles, no son negociables y, por ende, no tienen valor en el mercado. No obstante, y como en el caso de los activos económicos, sí pueden ser acumulados. La lucha

por derechos civiles es el mecanismo por el cual los activos políticos son acumulados. Los activos culturales pueden ser acumulados por personas y grupos a través de la educación, la migración, la organización social, el matrimonio y, sobre todo, a través de los derechos civiles.

Sigma es una economía sobrepoblada. En este sentido, se parece a la economía omega. Pero en sigma existe una desigualdad inicial en la dotación de activos económicos y sociales entre los individuos. Las características de la economía sigma pueden ser resumidas por el siguiente conjunto de proposiciones alfa:

$\alpha_\sigma(1)$. *Contexto institucional.* Reglas: los individuos participan en el proceso económico dotados de activos económicos y sociales. Existe propiedad privada de los activos económicos. Las personas pueden intercambiar los bienes sujetos a las normas del intercambio que rigen en el mercado, entre las cuales se incluye la norma de que los salarios nominales no pueden ser rebajados. También existen normas formales e informales de discriminación y segregación en el acceso a los activos culturales y políticos. Organizaciones: las organizaciones incluyen los hogares, las empresas capitalistas y el Gobierno. El Gobierno controla la oferta de dinero y de bienes públicos.

$\alpha_\sigma(2)$. *Condiciones iniciales.* Los individuos están dotados inicialmente con cantidades desiguales de activos económicos y sociales. Los más ricos tienen más de los activos sociales. Existen dos clases sociales: los capitalistas quienes concentran la propiedad del capital físico y los trabajadores, que no tienen capital físico. Entre los trabajadores existe un grupo con escasa dotación de activos sociales, la infraclase. Las dotaciones agregadas de capital y trabajo son tales que la economía es sobrepoblada.

$\alpha_\sigma(3)$. *Racionalidad de los agentes.* Los individuos actúan bajo la motivación del *homo economicus.* Los capitalistas bus-

can dos objetivos que están jerarquizados: primero, mantener su posición social; y, luego, maximizar las ganancias. En el mercado laboral, los trabajadores buscan el máximo salario al menor esfuerzo. Los capitalistas utilizan mecanismos para establecer la disciplina laboral. Los Gobiernos buscan maximizar el poder político.

$\alpha_{\sigma}(4)$. *Relaciones del mercado.* El sistema de mercado opera con mercados walrasianos y no-walrasianos. El mercado laboral es no-walrasiano. No existe mercado para los servicios de alquiler del capital físico; es decir, los propietarios no dan en arriendo su capital físico, ya que prefieren las utilidades a las rentas.

6.2. Un modelo estático de la economía sigma

Con el fin de deducir de esta teoría hipótesis empíricamente refutables, se debe establecer un modelo de la economía sigma. Hay tres grupos sociales: los capitalistas y dos clases de trabajadores con distintas dotaciones iniciales de capital humano. Los trabajadores con altos niveles de calificación serán llamados «trabajadores-y» y los de baja calificación serán llamados «trabajadores-z». Los capitalistas y los trabajadores-y tienen la misma dotación de capital humano y de activos políticos y culturales. Las empresas capitalistas emplean trabajadores-y para producir el bien B, la única mercancía producida en esta economía. Existe sobrepoblación de trabajadores-y, y eso implica la existencia de un sector de subsistencia.

Los trabajadores-z están dotados de menores cantidades de todos los activos —económicos y sociales—. En particular, son trabajadores con poco capital humano para la tecnología utilizada en el sector capitalista. Por lo tanto, su provisión de capital humano no es adecuada para el empleo asalariado. No son empleables. No son parte de la oferta de trabajo en el mercado laboral, es decir, están totalmente excluidos del mercado labo-

ral. Las empresas capitalistas no pueden obtener ganancias al emplearlos, ya que habría que invertir mucho en su capacitación, mientras que al mismo tiempo hay una abundante oferta de trabajadores-y. Es la lógica de la ganancia la que está detrás de la exclusión total de los trabajadores-z del mercado laboral. Los trabajadores-z constituyen una infraclase o subclase, pues son prescindibles desde el punto de vista de la racionalidad de los capitalistas.

Las empresas capitalistas procuran maximizar sus ganancias. Con el fin de lograr este objetivo, las empresas necesitan aplicar incentivos para extraer la óptima intensidad de esfuerzo de los trabajadores-y. Para hacer que la indisciplina laboral sea más costosa para los trabajadores, la tasa de salarios del mercado tiene que ser más alta que el costo de oportunidad de los trabajadores. En la economía sigma, este costo de oportunidad está dado por el ingreso que los trabajadores puedan lograr como productores auto-empleados en el sector de subsistencia. Se supone que la productividad marginal del trabajo en el sector de subsistencia está sujeta a rendimientos decrecientes. Cuanto mayor la cantidad de trabajadores auto-empleados, menor el ingreso marginal en este sector; alternativamente, cuanto menor la cantidad de mano de obra en el sector de subsistencia, mayor será el ingreso marginal. Por lo tanto, cuando las empresas quieren emplear más trabajadores-y, manteniendo la misma intensidad del trabajo, necesitan pagar mejores salarios porque el costo de oportunidad de los trabajadores aumenta.

El mercado laboral opera como en el caso de la economía omega. Las variables endógenas y exógenas son las mismas. A la tasa salarial del mercado habrá exceso de oferta laboral de los trabajadores-y. Este exceso de oferta no puede ser eliminado reduciendo los salarios nominales. El mercado laboral es no-walrasiano.

Al igual que en la economía omega, los trabajadores-y que son excluidos del mercado laboral pueden elegir, como segunda opción, entre desempleo y auto-empleo. Dada la curva

de productividad marginal decreciente en el sector de subsistencia, esta elección depende del salario esperado, que depende a su vez del salario de mercado. El salario del mercado debe ser mayor que el ingreso marginal del trabajador en el sector de subsistencia como mecanismo que asegura la disciplina laboral.

Los trabajadores-z se auto-emplean en pequeñas unidades de producción, donde producen el bien B. En consecuencia, en la economía sigma existen dos sectores de subsistencia. En ambos sectores se produce el único bien que existe en la economía. No hay, por lo tanto, necesidad de intercambio entre estos dos sectores. En el sector de subsistencia-z, la tecnología es tradicional y el nivel de la productividad laboral es menor comparada al sector de subsistencia-y.

6.3. Equilibrio general

El modelo de la economía sigma tiene tres sectores: el sector capitalista y dos sectores de subsistencia: el de los trabajadores-z y el de los trabajadores-y. El sector capitalista y el sector de subsistencia-y operan e interactúan como en la economía omega que se presentó en el capítulo anterior. Así, existen cuatro mercados en la economía sigma: el de la mercancía, el del dinero, el laboral y el de la divisa. En esta subeconomía omega existe un subsistema de relaciones que da lugar a un equilibrio general.

En el sector de subsistencia-z, los trabajadores buscan maximizar el producto total, sujeto a la restricción de su cantidad de mano de obra y la tecnología. Debido a que no existe intercambio entre este sector y el resto de la economía, el equilibrio en este sector es simple de establecer y es separable.

En suma, el equilibrio general en la economía sigma se establece de manera secuencial y separada. Primero se determina el equilibrio en el mercado laboral y el mercado de dinero, donde se determinan el tipo de cambio y la cantidad de empleo asala-

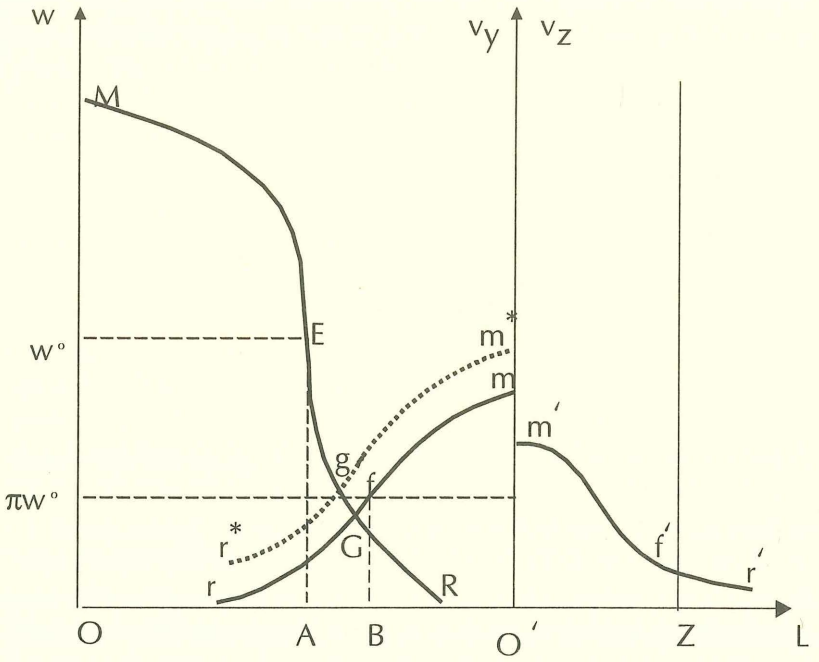
riado. Con estas variables ya conocidas, la producción y distribución en el sector capitalista queda determinada. También queda determinada la cantidad de exceso de oferta laboral de los trabajadores-y. En segundo lugar, con las variables determinadas en el sector capitalista, que es donde opera el sistema de mercado, se determina de manera simultánea el desempleo y el auto-empleo de los trabajadores-y. El ingreso total y el ingreso medio en el sector de subsistencia también quedan determinados. En tercer lugar, el ingreso total y el ingreso medio en el sector de subsistencia-z se determina independientemente del resto de la economía.

El equilibrio general en la economía sigma se representa en la Figura 6.1. La cantidad de trabajadores-y se mide por el segmento OO'. Hay dos planos en este gráfico. En el primer eje vertical se mide el producto por trabajador en el sector capitalista. La curva MER representa la curva de la productividad marginal del trabajo en las firmas capitalistas, que es igual a la curva de demanda de trabajo. En el segundo eje vertical se mide el producto por trabajador en el sector de subsistencia-y. La curva *mr* representa la curva de la productividad marginal del trabajo (medida desde el origen O' y hacia el origen O), y la curva m^*r^* es la correspondiente curva de extracción de esfuerzo que asegura la posición de la curva de productividad del trabajo al nivel dado por MER.

La solución de equilibrio general implica en el mercado laboral un salario real igual a w^o y un nivel de empleo asalariado igual a OA. El exceso de oferta laboral resultante es igual a AO'. Este exceso se distribuye, dado el valor del salario real, entre AB al desempleo y BO' al auto-empleo.

Para el sector de subsistencia-z, el producto por trabajador se mide también en el segundo eje vertical y la cantidad de trabajadores en el eje horizontal, en el plano que tiene por origen O'. La curva $m'r'$ representa la curva de la productividad marginal del trabajo en este sector. La cantidad de trabajadores es igual al segmento O'Z en el eje horizontal.

Figura 6.1
Equilibrio general en la economía sigma



El ingreso generado en el sector capitalista es igual al área debajo del segmento ME de la curva de la productividad marginal del trabajo MER. El ingreso generado en el sector de subsistencia-y es igual al área debajo del segmento *mf* de la curva de la productividad marginal del trabajo *mr*. El ingreso generado en el sector de subsistencia-z es igual al área debajo del segmento *m'f'* de la curva del producto marginal del trabajo *m'r'*. El ingreso nacional es igual, entonces, a la suma de estas tres áreas. Dado que la economía produce un único bien, esta agregación es dimensionalmente correcta.

El ingreso nacional (Y) de equilibrio y su distribución pueden ser representados por el siguiente sistema de identidades:

$$
\begin{aligned}
Y &= P + W + V_y + V_z \\
&= P + w D_y + v_y L_y + v_z L_z \qquad (6.1) \\
S_{hy} &= D_y + L_y + U_y \\
S_{hz} &= L_z
\end{aligned}
$$

La primera ecuación indica que el ingreso nacional se compone de ganancias, masa salarial, el ingreso total en el sector de subsistencia de los trabajadores-y y el total de ingresos en el sector de subsistencia de los trabajadores-z. La segunda ecuación muestra esa misma distribución pero en términos de ingresos promedios y cantidad de trabajadores. La tercera ecuación señala la asignación de la oferta laboral de los trabajadores-y al empleo asalariado, al auto-empleo y al desempleo. La última ecuación indica que la cantidad de trabajadores-z es asignada solo al auto-empleo.

El equilibrio general incluye la condición de que el ingreso medio en el auto-empleo tiene que ser menor que el salario de mercado entre los trabajadores-y ($v_y < w$). Las dotaciones iniciales en activos dan lugar a la siguiente relación en los ingresos medios de los trabajadores: $v_z < v_y < w$. Los trabajadores-z constituyen el grupo social más pobre en la sociedad sigma.

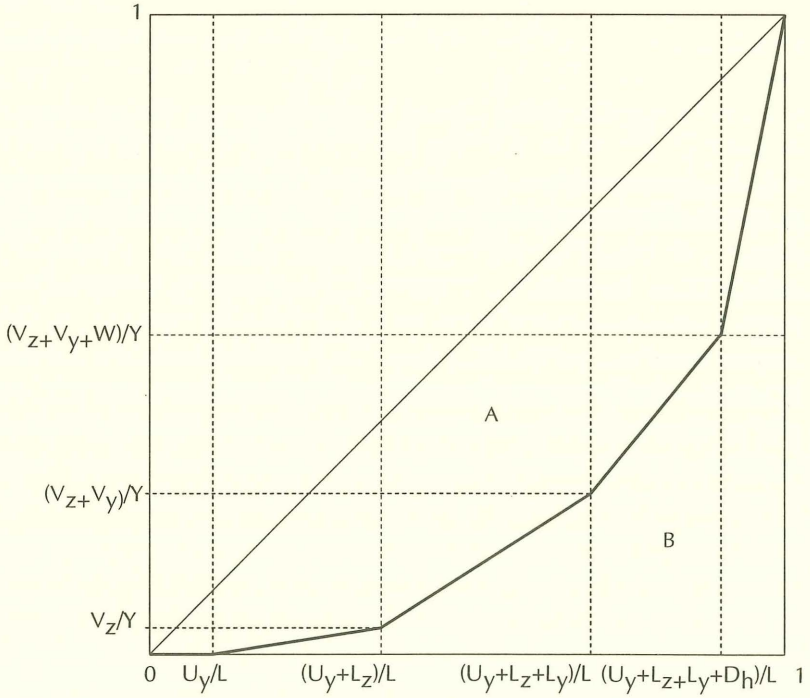
La sociedad sigma muestra una dispersión de ingresos importante. Existe desigualdad no solo entre capitalistas y trabajadores sino, también, entre trabajadores. Además, no todos los trabajadores-y obtienen el mismo ingreso; tampoco todos los trabajadores-z reciben el mismo ingreso. Comparado a una sociedad omega, con iguales dotaciones de capital y tecnología, la economía sigma es más desigual. Aquí se ha introducido un grupo de trabajadores sujetos a la exclusión social que tienen ingresos más bajos que el resto de la sociedad.

La concentración del ingreso que resulta del proceso económico en la sociedad sigma se puede representar por una curva de Lorenz. Esta curva se muestra en la Figura 6.2. Tiene cinco segmentos que corresponden a los distintos grupos sociales y a sus posiciones relativas en la distribución del ingreso nacional en la situación de equilibrio de la economía. Ciertamente, la

curva de Lorenz en la sociedad sigma muestra una sociedad más compleja que aquellas que aparecieron en las curvas de Lorenz presentadas anteriormente para las sociedades épsilon y omega.

En la economía sigma, las variables exógenas incluyen el *stock* de capital y el conocimiento tecnológico, así como la desigualdad inicial en la dotación individual de activos. Para el corto plazo, la oferta de dinero y los *shocks* externos son también exógenos. Entre las variables endógenas se encuentra el nivel del ingreso nacional y su distribución. Si las variables exógenas no se modifican, los valores de equilibrio del ingreso

Figura 6.2.
Curva de Lorenz sobre la distribución de ingresos en la economía sigma



nacional y su distribución se repetirían período tras período. Cambios en las variables exógenas modificarían los valores de equilibrio del nivel del ingreso nacional en direcciones particulares, pero no modificarían la desigualdad en la distribución del ingreso. Así, las predicciones empíricas pueden ser derivadas lógicamente de este modelo teórico.

6.4. Proposiciones beta

La estructura de la economía sigma se compone de una combinación de la economía omega y del sector de subsistencia-z, donde este sector es separable del resto de la economía. La solución de equilibrio general de la economía omega es todo lo que se necesita resolver para tener el equilibrio general de la economía sigma. Se puede entonces utilizar aquí los resultados obtenidos en el capítulo 5. El equilibrio general en la sociedad omega se obtiene de manera secuencial, primero el sistema de mercado resuelve la producción y distribución en el sector capitalista y, solo una vez conocido estos valores de equilibrio, se resuelve el equilibrio en el sector de subsistencia. En la economía sigma se puede utilizar este mismo principio, donde solo hay que precisar que el sector de subsistencia se refiere al sector de subsistencia-y.

El equilibrio en los mercados se puede establecer en la economía sigma también utilizando el sistema de curvas HH-MM, que constituye el mismo sistema que se utilizó en la economía omega y, antes, en la economía épsilon. El resultado es que el sector capitalista opera de manera similar en los tres tipos de economías, a pesar de que los mecanismos de extracción de esfuerzos son diferentes. Por lo tanto, es suficiente aplicar la estática comparativa al sistema HH-MM para obtener las proposiciones beta de toda la economía sigma.

Los resultados se presentan en el Cuadro 6.1. El primer grupo de variables endógenas se refiere al sector capitalista, el segun-

do grupo al sector de subsistencia-y y el tercero a las variables distributivas. Las variables exógenas y sus efectos sobre las variables endógenas son las mismas que se obtuvieron en la economía omega. Debido a que existe un nivel de empleo máximo en el sector capitalista, estos efectos corresponden al caso en el que el nuevo equilibrio del empleo asalariado se da por debajo de ese nivel máximo ($D_{hy} < L_y^*$). Si la cantidad de dinero aumentara cuando el equilibrio del empleo asalariado es igual al del nivel máximo ($D_{hy} = L_y^*$), el efecto sería aumentar el nivel de precios solamente. En sigma, la neutralidad del dinero aparece cuando se

Cuadro 6.1
Matriz de proposiciones beta de un modelo sigma estático ($D_h<L^*$)

Variables endógenas	Variables exógenas		
	S_m	z^*	K_b
P_e	+	−	−
D_h	+	+	+
Q_b	+	+	+
w	−	+	+
W	?	+	+
P	+	+	+
P_b	+	−	−
P_h	0	0	0
L_y	+	−	−
v_y	−	+	+
V_y	+	−	−
U_y	−	?	?
Y	+	+	+
P/Y	?	?	?
W/Y	?	?	?
V_y/Y	?	−	−
V_z/Y	−	−	−
w/v_y	?	?	?
w/v_z	−	+	+

llega al umbral del empleo máximo de tipo asalariado, que a su vez depende del umbral del salario de eficiencia w^*.

Es evidente que las variables exógenas no tienen efecto algunos sobre el sector de subsistencia-z. La acumulación de capital en el sector capitalista, por ejemplo, no tiene efecto sobre este sector. No tiene efecto sobre su nivel de empleo, pues estos trabajadores están excluidos del mercado laboral; tampoco afecta su curva de productividad, pues estos trabajadores no tienen el capital humano necesario para adoptar las nuevas tecnologías. Las diferencias en productividad entre los sectores se mantienen. El efecto de «chorreo» *(trickle down)* del crecimiento en el sector capitalista sobre este sector es nulo.

El ingreso y el empleo en el sector de subsitencia-z están exógenamente determinados. Las posibles interacciones entre este sector y el resto de la economía pueden resumirse como sigue:

(a) La acumulación de capital junto con el cambio tecnológico en el sector capitalista (un desplazamiento hacia arriba de la curva de demanda de trabajo) no tendría efecto alguno sobre el sector de subsistencia-z. El proceso de acumulación de capital puede finalmente eliminar el sector de subsistencia-y, pero el sector de subsistencia-z permanecerá inalterado.
(b) No hay convergencia en la productividad de mano de obra entre el sector capitalista y el sector de subsistencia-z.
(c) Aumentos en la población-z reducirían el valor de la productividad marginal de la mano de obra en el mismo sector, y eso reduciría el ingreso medio.
(d) Aumentos en la productividad en el sector de subsistencia-z no tendrían efecto sobre el resto de la economía.

Estas proposiciones beta señalan que, en la economía sigma estática —en la que la inversión en capital físico es exógena—, el sector de subsistencia-z no afecta al sector capitalista ni es afectada por ese sector. Si bien este modelo de sigma predice una dualidad persistente entre el sector capitalista y el sector de

subsistencia-z, no predice lo mismo para la relación entre el sector capitalista y el sector de subsistencia-y. En este caso, el crecimiento en el sector capitalista afecta (y positivamente) al sector de subsistencia-y.

El último grupo de variables endógenas en el Cuadro 6.1 se refiere a la distribución del ingreso. El efecto de las variables exógenas sobre la desigualdad global es ambiguo, pues sus efectos sobre los componentes de la distribución agregada son indeterminados. Debido a que el ingreso medio en el sector de subsistencia-z es fijo, cualquier aumento en los salarios (debido a un aumento en el *stock* de capital o en los términos de intercambio) generará una pérdida relativa para ese grupo. Pero, al mismo tiempo, la distribución en la subeconomía omega entre el ingreso en el sector de subsistencia, la masa salarial y la ganancia es indeterminada, como se mostró en el capítulo anterior, de modo que no se puede concluir sobre el cambio en la desigualdad global.

6.5. La naturaleza del equilibrio en la sociedad sigma

El equilibrio general estático que se ha establecido en la sociedad sigma implica que los valores de equilibrio de las variables endógenas del sistema se repetirán periodo tras periodo en tanto los valores de las variables exógenas permanezcan fijas. ¿Existen mecanismos adicionales que puedan cambiar esta situación?

Exclusión del mercado de crédito y de seguros

El mercado laboral genera desigualdad entre los trabajadores-y. Los trabajadores-y que se encuentran excluidos del mercado laboral tienen las mismas calificaciones de aquellos que se encuentran incluidos; por lo tanto, ¿por qué los primeros tendrían que permanecer relativamente pobres? Ellos podrían establecer una nueva empresa o incrementar la productividad en las unidades del sector de subsistencia. ¿Por qué no lo hacen?

Debido a que los propietarios no dan en arriendo su capital, los trabajadores no pueden establecer empresas mediante el alquiler de capital. Tienen que comprar capital. Estos trabajadores no tienen suficientes ahorros para financiar la acumulación de capital. Requieren crédito bancario.

Sin embargo, la lógica de los bancos es tal, que los trabajadores son excluidos de este mercado. Sobre la lógica de los bancos ya se presentó la teoría respectiva (Capítulo 4). Es suficiente señalar ahora que los bancos buscan un doble objetivo: maximizar las utilidades y minimizar los riesgos. Debido a que el costo unitario por dólar se reduce según la magnitud del préstamo, los bancos prefieren otorgar préstamos grandes; más aún, dado que la información es incompleta, los bancos minimizan los riesgos pidiendo garantías cuyo valor depende positivamente del volumen del préstamo. Por esta razón, los bancos establecen umbrales en la magnitud de préstamo y riqueza, con lo que se determina quiénes serán sus clientes. La lógica de los bancos excluirá a los potenciales prestatarios cuyo capital sea inferior al valor del umbral o a aquellos que simplemente no lo posean. El mercado del crédito es no-walrasiano.

También existe el riesgo de destrucción de los acervos de capital. Este riesgo puede ser asegurado por medio del mercado de seguros, un instrumento para diluir el riesgo. Sin embargo, el mercado de seguros tampoco es walrasiano. Las razones son similares a las del mercado de crédito. Los costos de operación para asegurar empresas pequeñas son demasiado altos. Las empresas aseguradoras prefieren hacer negocios con grandes empresas. Por consiguiente, los productores en el sector de subsistencia están sujetos a riesgos individuales no asegurables. Son, por lo tanto, vulnerables a los *shocks* negativos a su dotación de capital.

En resumen, como resultado de la racionalidad económica de los bancos y de las empresas aseguradoras, y de los costos de la información, los trabajadores-y que son excluidos del mercado laboral, también son excluidos de los mercados del crédito y de los seguros. No pueden, en consecuencia, escapar

de su pobreza relativa. El auto-empleo y el desempleo son sus únicas posibilidades, pero son opciones en una situación de segunda opción. Los trabajadores-z también son excluidos de los mercados del crédito y de los seguros. Los factores que explican la exclusión de los trabajadores-y de estos mercados operan con mayor intensidad para los trabajadores-z. Estos últimos tampoco pueden escapar de su pobreza.

La economía sigma tendrá un segmento de la fuerza laboral que no es asalariado. Este segmento incluye parte de los trabajadores-y y a todos los trabajadores-z. Ellos son excluidos del mercado laboral y del mercado de crédito, que son los dos mecanismos a través de los cuales se puede explotar a los trabajadores, como lo ha demostrado Roemer (1982). Los asalariados —aquellos trabajadores-y que se encuentran totalmente integrados en el sistema capitalista— son los más ricos entre todos los trabajadores. Los excluidos son más pobres; y, a causa de que el grado de exclusión es más alto para los trabajadores-z, ellos constituyen el grupo más pobre. El muy conocido aforismo de Joan Robinson —«la única cosa que es peor que ser explotado, es no ser explotado»— se aplica con mucha fuerza a la sociedad sigma.

Las reglas del intercambio en microsociedades

Debido a la exclusión del mercado de seguros que sufren los trabajadores-z, ellos buscan formas individuales para auto-asegurarse contra los riesgos y, también, formas colectivas de esparcir los riesgos. Y debido a su exclusión del mercado de crédito, ellos buscan préstamos de su amigos. De este modo, las redes sociales son creadas como un mecanismo que provee protección social y oportunidades. Como resultado de su exclusión de los mercados financieros y de seguros, los trabajadores-z forman redes en la forma de comunidades o microsociedades.

Dentro de una microsociedad cerrada, el intercambio económico no puede tomar la forma del de mercado. En una sociedad tan pequeña, la interdependencia individual es muy alta y

predomina el intercambio múltiple. De ahí que las relaciones de intercambio no puedan ser impersonales. La microsociedad está compuesta de redes creadas para resolver el problema de riesgos individuales, un problema que no puede ser solucionado por el sistema de mercado. En las microsociedades, la racionalidad del interés propio no puede ser la única motivación de las personas, ya que las consecuencias económicas pueden ser negativas para los demás. Los comportamientos individualistas basados en este razonamiento pueden generar pérdidas económicas para el resto de los miembros de la comunidad. Entonces, el individuo podría sufrir una sanción social por ese comportamiento. Dado el hecho de que ese intercambio múltiple prevalece en la comunidad, la persona podría ser excluida de algunos de los tipos de intercambio y también sufrir pérdidas económicas.

Pero, ¿cuáles son las reglas del intercambio fuera del mercado? El intercambio continúa siendo voluntario y basado en la lógica del interés propio, pero los individuos están restringidos por las normas sociales de la red social, en adición a las restricciones que les impone su limitada dotación de recursos. Estas normas incluyen reciprocidad y redistribución. «Hoy por ti, mañana por mí» es un principio de intercambio. El balance económico del intercambio se obtiene en el largo plazo, no en cada transacción. De ahí que la ley del precio único para cada bien pueda no tener vigencia. Los beneficios en el intercambio dentro de la red son tan importantes (como la protección social) que debilitan el incentivo de llevar a cabo el arbitraje en los precios y llegar, así, a un precio único para cada bien en la microsociedad. Además, el dinero no puede comprar todo dentro de estas redes sociales. Estas son diferencias importantes en relación con las normas del intercambio del mercado.

La red social construida por los pobres es una estrategia de supervivencia, no es una estrategia de desarrollo. A cada trabajador le permite reducir el riesgo. Es un mecanismo que provee una red de seguridad. Ningún individuo se convertirá en indigente por un *shock* negativo sobre sus dotaciones de recursos.

Sin embargo, al mismo tiempo, ningún individuo puede escapar fácilmente de la red —esto es, de la pobreza— aunque se le presente una oportunidad externa. Las reglas de la reciprocidad y la redistribución establecen límites para sus opciones. Esta es la otra cara de la moneda de la protección social. El intercambio fuera del mercado protege a los individuos contra los riesgos, pero los condena a compartir ingresos extraordinarios con todo el grupo. Por contraste, el intercambio de mercado no protege a los individuos colectivamente contra el riesgo, pero les permite la posibilidad de escapar de la pobreza individualmente.

El contrato social en una microsociedad es rawlsiano. Rawls (1971), en efecto, utilizó una comunidad pesquera para ilustrar su teoría del contrato social basado en la equidad distributiva.

Si las microsociedades se abren al sistema económico mayor, como la economía capitalista, las reglas del intercambio serán duales. Los individuos tendrán acceso al intercambio en el mercado. Por lo tanto, las normas sociales dentro de la comunidad no serán tan rígidas como en el caso anterior. Mientras más desarrollados sean los mercados, más débiles son las restricciones determinadas por las normas sociales de la comunidad. En la economía sigma, los trabajadores-z viven segregados en microsociedades, en las que las reglas de intercambio fuera del mercado predominan en el conjunto de todas las transacciones.

La utilización de las redes sociales y del intercambio fuera del mercado no es, sin embargo, una exclusividad de los trabajadores-z. Todos los individuos siguen esta estrategia porque hay retornos económicos a invertir en redes sociales. Ciertamente, este retorno económico es mayor para los individuos que están dotados con mayores cantidades de activos económicos y sociales. Se puede esperar que los capitalistas obtengan los más altos retornos de sus redes sociales debido, entre otras cosas, a que estas redes son parte de sus estrategias tanto de sobrevivencia (como capitalistas) como, también, de desarrollo.

Diferencias persistentes en la productividad

Dadas las diferencias en tecnología entre el sector capitalista y el sector de subsistencia-y, este sector debería ser capaz de adoptar esa tecnología y elevar su curva de productividad. Aún más, debería crecer más rápido que el sector capitalista, pues, dado su retraso tecnológico, hay mucho espacio para el crecimiento de la productividad en el sector de subsistencia-y. Para la adopción de nuevas tecnologías, el trabajador-y no podría quedar limitado por la falta de capital humano necesario par adoptar las innovaciones, puesto que ellos tienen el mismo capital humano que el resto de trabajadores en el sector capitalista. El factor que limita este desarrollo del sector de subsistencia-y es el funcionamiento del mercado de crédito y del de seguros. El proceso de adopción tecnológica requiere financiamiento. Sin embargo, los trabajadores-y carecen de ahorros y, como ya vimos, están excluidos de los mercados de crédito y de seguros.

Los trabajadores-z también permanecen pobres, aunque su existencia en una sociedad capitalista les debería dar oportunidad para adoptar innovaciones tecnológicas y así incrementar su productividad. En realidad, el sector-z también debería poder crecer más rápido que el sector capitalista, en el que las innovaciones solo ocurren intermitentemente. En este caso, la limitación proviene también de la dotación de capital humano de los trabajadores-z, que es muy baja para la actual etapa de desarrollo tecnológico. Y aunque hubiera un subconjunto de nuevas tecnologías que ellos pudieran adoptar, no lo podrían hacer ya que necesitarían financiamiento externo y medios para esparcir el riesgo. Pero, al igual que los trabajadores-y, ellos también son excluidos de los mercados de crédito y de seguros.

Diferencias en la acumulación de capital humano

La teoría sigma supone que los trabajadores-z están dotados inicialmente de un bajo nivel de capital humano. Pero, ¿por qué

los trabajadores-z no pueden acumular capital humano? La respuesta a esta pregunta no puede ser otro supuesto sino una derivación lógica de los supuestos de la teoría, es decir, este resultado debe ser endógeno.

Existen varios mecanismos. Está, en primer lugar, su pobreza. Carecen de la capacidad financiera que se requiere para acumular capital físico y humano. En segundo lugar, están excluidos de los mercados de crédito y de seguros.

El tercer factor tiene que ver con la exclusión social. Debido a que son ciudadanos de segunda categoría, no tienen acceso a derechos económicos (en cantidad y calidad) en la forma de bienes públicos, como la educación, la salud y la protección social. Este es el problema de la exclusión política. También viven segregados, y eso hace que el aprendizaje de habilidades para la tecnología moderna sea muy costoso. Este es el problema de la exclusión cultural.

Las teorías psicológicas sugieren que la inteligencia es resultado tanto de la herencia genética de los individuos como de su medio social. La gente segregada de un ambiente social estimulante —asociado a la vida moderna— no se beneficia del efecto positivo que tiene este ambiente sobre el desarrollo cerebral. En una sociedad sigma, la distribución de la inteligencia entre la población es entonces endógena. La consecuencia de este resultado para la acumulación de capital humano es que el costo de la acumulación sería mayor para la población socialmente excluida.

Derechos, acción colectiva y la lógica de los Gobiernos

Sigma es una sociedad democrática. Entonces, ¿por qué los trabajadores-z no pueden eliminar la exclusión por medio de una *demanda por* derechos sociales? Esto requeriría una acción colectiva, pues se trata de producir un bien público.

El problema olsoniano sería la primera limitación a mencionar. Olson intentó responder esta pregunta: ¿por qué hay tan poca acción de clase en el mundo real? Su teoría es que la

acción colectiva no ocurrirá si los individuos que integran el grupo actúan guiados por la lógica del interés propio (Olson, 1965). En las microsociedades, sin embargo, los «gorreros» sufrirían sanciones sociales; luego, el problema olsoniano no aparecería. Los limites de la acción colectiva provienen, entonces, del problema de la exclusión. Los trabajadores-z están demasiado mal dotados de activos políticos y culturales para participar ampliamente en el juego democrático.

Pero también la pobreza misma juega en contra de la acción colectiva. Los trabajadores-z son muy pobres para dedicar sus magros ingresos a financiar una actividad colectiva cuyo producto es un incierto bien público. Son muy pobres para gastar parte de sus ingresos en este bien. Este es el problema maslowniano. El psicólogo Maslow (1970) propuso la teoría de que los individuos buscan satisfacer sus necesidades y dichas necesidades están jeraquizadas en necesidades primarias y secundarias. Solo después de satisfacer las primarias se pasa a las necesidades secundarias. Bienes públicos inciertos no se encuentran entre las necesidades primarias —que son las que buscan satisfacer los trabajadores-z—. Los trabajadores-z son muy pobres para gastar en la obtención de derechos.

Cualquiera que fuesen los méritos empíricos de las teorías de Olson y Maslow, en la sociedad sigma, el factor más importante tiene que ver con la exclusión social misma. Debido a la desventaja acumulativa que los trabajadores-z tienen en la sociedad, su capacidad para auto-organizarse y demandar acceso a los derechos es limitada. En este caso, la acción colectiva queda limitada por el problema de la exclusión social. La limitación está en la falta de voz, debida a la exclusión política y cultural. En el lado de la demanda por derechos, la exclusión social crea su propio círculo vicioso.

En el lado de la oferta, los factores tampoco son favorables. En la sociedad sigma, ¿tendrían los Gobiernos incentivos para desarrollar derechos? Esta pregunta nos lleva a discutir la lógica de los Gobiernos. El supuesto usual es que los Gobiernos buscan

maximizar el poder político. Esta proposición se puede poner en un modelo sencillo que sostiene que los Gobiernos buscan maximizar votos y lo hacen sujetos a su restricción presupuestaria. La predicción de este modelo es que los Gobiernos gastarán el presupuesto público tratando de obtener la máxima rentabilidad política. Si los Gobiernos gastan en cumplir con la provisión de derechos ya establecidos, la rentabilidad política de ese gasto será nula o muy baja comparada a un gasto discrecional que sí genera clientelismo entre la población. Los Gobiernos prefieren gastar de manera discrecional. Su lógica no los lleva a crear derechos.

En suma, en la sociedad sigma, el equilibrio se da con exclusión social. La transformación de una sociedad heterogénea y jerarquizada, en una sociedad homogénea —la transformación de los trabajadores-z en trabajadores-y— no podría ocurrir de manera espontánea ni automática. Esta transformación no puede ser endógena.

Por lo tanto, los trabajadores-z constituirán el núcleo duro de la pobreza En la sociedad sigma, los grupos más pobres no solo son personas con ingresos bajos en relación con el resto de la sociedad, sino que son personas étnicamente diferentes al resto de la sociedad. Los más pobres son los más excluidos, y viceversa. Esta es la característica de la pobreza en una sociedad sigma.

Tres mercados juegan un rol crucial en la reproducción de la desigualdad en la economía sigma: el laboral, el de crédito y el de seguros. Estos mercados pueden ser llamados *mercados básicos*. Dada la condición inicial de desigual distribución de los activos económicos, políticos y culturales en la sociedad sigma, la desigualdad será reproducida en el proceso de acumulación de capital en el sector capitalista. Un segmento de los trabajadores-y será excluido de los mercados básicos. Los trabajadores-z se quedan detrás en el proceso de crecimiento económico, ya que son excluidos no solo de los mercados básicos sino, también, de los derechos políticos y culturales. Esta última exclu-

sión —la que denominamos «exclusión social» para diferenciarla de la exclusión económica— pone a los trabajadores-z aún más restricciones para la acumulación de capital, en particular para la acumulación de capital humano. Lo que es particular en la sociedad sigma —como una democracia capitalista— es la exclusión social, pues la exclusión económica existe en las tres sociedades estudiadas.

Aunque la distribución de activos es individual, el mecanismo de exclusión no opera sobre bases individuales. Los individuos tienen afiliación a grupos étnicos. La exclusión es sistemática para grupos étnicos determinados. El concepto de exclusión se refiere al grupo y no pretende explicar casos individuales de éxito económico.

6.6. Consistencia empírica: los países latinoamericanos

Debido a la escasez de datos sobre la producción y la distribución en el tercer mundo, la teoría sigma se pondrá a la prueba empírica en un contexto más limitado. Se utilizará fragmentos de evidencia sobre los países latinoamericanos para desafiar la hipótesis de que los países con una amplia proporción de población indígena funcionan como si fueran economías sigma, donde la población indígena representa trabajadores-z. Parece indiscutible el hecho de que esta población está dotada de la menor cantidad de activos económicos, políticos y culturales (según la valorización de la cultura predominante). La proporción de población indígena se utilizará como indicador para determinar a los «países indígenas».

Países en donde la proporción de poblaciones indígenas es más alta serán más desiguales

En América Latina existe una diversidad de países en términos de etnicidad. Se pueden distinguir dos grupos extremos.

El grupo más homogéneo incluye a Argentina, Uruguay, Venezuela y Costa Rica. El grupo más heterogéneo incluye a México, Guatemala, Colombia, Ecuador, Perú y Bolivia. La conquista por España puede ser considerada como el impacto fundacional. Cinco siglos después de la Conquista, todavía hoy en día las poblaciones indígenas constituyen una significativa proporción de la población total de estos países. Los miembros de estas poblaciones aún mantienen parte de su cultura ancestral y viven en su mayoría en comunidades campesinas rurales. El Brasil puede ser incluido en este segundo grupo. La presencia de negros —que fueron traídos como esclavos desde el África—, sumado al hecho de la presencia de comunidades indígenas, ha creado una sociedad que es multiétnica, situación que también se da en el caso de los restantes países del grupo.

Estudios realizados en los años setenta sobre la distribución del ingreso nacional —no solo sobre el ingreso del trabajo, como se muestra en los estudios más recientes— mostraron que la desigualdad de ingresos en el segundo grupo era consistentemente más alta que en el primer grupo durante los cincuenta y sesenta. Como se muestra en Cuadro 6.2, los coeficientes Gini promedios para Costa Rica y Argentina fueron 0,37 y 0,43; mientras que para el Perú, Brasil, Colombia y México fueron del rango de los 0,54 a 0,62; y la cifra para Chile fue 0,50. Chile es un caso intermedio en términos de heterogeneidad social.

En el Cuadro 6.2, también se muestra la más reciente y completa información sobre la distribución del ingreso nacional, sobre la base de cinco o más observaciones en el período 1950-1993. Este conjunto de datos se presenta para solamente cinco países. Esta información estadística también confirma la alta y positiva correlación que parece existir entre heterogeneidad social y desigualdad. El bajo valor del coeficiente de variabilidad de Pearson para cada país sugiere que la distribución de ingresos no ha variado mucho dentro de cada país. Esta persistencia de largo plazo tanto en el grado de desigualdad, así como en la

Cuadro 6.2
América Latina: desigualdad en ocho países, en años seleccionados entre 1950-1993

	Weisskoff-Figueroa 1950-1970		Altimir 1950-1993		
País	Nº de observaciones (años)	Coeficiente Gini Valor medio	Nº de observaciones (años)	Coeficiente Gini Valor medio	C.V. de Pearson (%)
Brasil	2	0,58	7	0,62	4,0
Perú	1	0,62	-	-	-
México	3	0,54	8	0,55	4,0
Colombia	1	0,58	6	0,50	6,7
Chile	1	0,50	-	-	-
Argentina	3	0,43	-	-	-
Venezuela	-	-	5	0,42	9,8
Costa Rica	1	0,37	9	0,42	4,1

- No disponible
Fuente: Weisskoff y Figueroa (1976: 91); las estimaciones de Óscar Altimir aparecen en Thorp (1998: 352, Tabla de Apéndice VIII.1).

categoría de desigualdad por países, puede ser sorprendente, pero es consistente con la predicción de la teoría sigma.

Considerando los dos grupos de países mencionados más arriba, el segundo grupo nació al capitalismo socialmente más heterogéneo y más desigual que el primero. Este resultado empírico es consistente con la teoría sigma. El grado de desigualdad de ingresos parece ser una característica estructural de cada país. El naturalista alemán Alexander von Humboldt escribió, en su *Viaje*, que la desigualdad económica y social fue la característica que más le impresionó en su visita a la región ahora conocida como América Latina. Ese trabajo fue publicado alrededor de 1850. Ciento cincuenta años después, la desigualdad todavía es una característica central de la región.

Se puede demostrar que el mayor grado de desigualdad en el segundo grupo de países se debe, en gran parte, a la alta

proporción de su población indígena. Si recalculáramos el ingreso nacional omitiendo la participación de la población indígena, el resultado sería un nivel ligeramente más bajo del ingreso nacional, pero sería una sociedad significativamente más igualitaria. La razón es simple: una gran cantidad de personas con un pequeño monto de los ingresos serían separados del ingreso nacional original.

Tomemos el caso del Perú. En la mitad de los setenta, el decil más alto de la población recibió el 50% del ingreso nacional, mientras que el tercio inferior solo recibió el 5% (Webb 1977). La población indígena constituía la gran mayoría de este grupo bajo. Con el recálculo, la razón entre el decil más alto y el tercio inferior se reduce de 10 veces a 4 veces. El ingreso promedio se incrementa en un 40%.

Las poblaciones indígenas se encuentran en la base de la pirámide de ingresos en todos los países de la región. En los años ochenta, la incidencia de la pobreza en las poblaciones indígenas de Bolivia era de 64% (contra 48% entre los no indígenas); en Guatemala, de 87% (contra 54%); en México, 81% (contra 18%); y en el Perú, 79% (contra 50%) (Psacharopoulos y Patrinos 1994).

El núcleo duro de la pobreza se encuentra en las poblaciones indígenas, tal como predice la teoría sigma. Y como están dotadas de los más bajos montos de activos económicos, políticos y culturales, las poblaciones indígenas parecen poseer los atributos de los trabajadores-z.

En el caso del Brasil, a fines de la década de 1990, las estadísticas oficiales mostraron que 44% de la población total fue clasificada como población negra. Pero, en el quintil inferior de la distribución de ingresos, esta proporción era de 69% (*Jornal de Commercio,* Recife, noviembre 21, 1999, p. 16). Entonces, en Brasil, la población negra parece conformar el concepto de población-z. Joaquim Nabuco, un abolicionista brasileño, argumentó —hacia los años de 1880— que la abolición de la esclavitud debería estar acompañada de una reforma agraria. Sin

tierra —dijo— los manumitidos continuarían trabajando para los mismos dueños de la tierra y sus derechos de ciudadanía no estarían garantizados. En términos de la teoría sigma, resulta evidente que Nabuco estaba proponiendo un *shock* refundacional para Brasil, que no se llevó a cabo.

La persistencia de la economía campesina en los países indígenas

Como lo predice la teoría sigma, el crecimiento económico en el sector capitalista no ha eliminado la economía campesina en América Latina. La economía campesina la componen en su mayoría las poblaciones indígenas. Estas poblaciones han estado sujetas a los mecanismos de exclusión durante el proceso de crecimiento económico que ha tenido lugar en el sector capitalista. Esa exclusión es lo que podría explicar el hecho de que la expansión del capitalismo en Latinoamérica no haya logrado producir la gran transformación del campesino en asalariado, como sí lo hizo en Europa.[1]

[1] Al final de los setenta, el distinguido historiador Eric Hobsbawm dio una conferencia en mi universidad y la inició haciéndole la siguiente pregunta a la audiencia: «¿Por qué todavía observamos campesinado en los Andes peruanos?». Nadie respondió. Entonces nos explicó el caso de Escocia. En el siglo XVIII, las montañas de Escocia también estaban pobladas de campesinos. Menos de un siglo de desarrollo capitalista fue suficiente para transformar al campesinado en fuerza de trabajo asalariada. Hoy día, esas montañas están vacías, concluyó. Muchos años después, creo que ya tengo la respuesta a su pregunta. Perú nació como sociedad sigma (del mismo modo que todos los países latinoamericanos con civilizaciones antiguas importantes); y, por lo tanto, el campesinado ha estado siempre sujeto a los mecanismos de exclusión social. Este no fue el caso en Europa. Los antiguos países europeos se asemejan a la sociedad omega, esto es, sobrepoblada, pero sin trabajadores-z. Estos países se desarrollaron mucho más fácilmente para convertirse en sociedades épsilon. Para transformarse en sociedades épsilon, las sociedades sigma deben convertirse primero en sociedades omega, pero la exclusión social impone barreras.

El hecho de que en los países no indígenas haya campesinos puede ser el resultado de un proceso diferente: tales países comenzaron sobrepoblados pero no socialmente heterogéneos. En cualquier caso, esto no refuta a la teoría sigma. La teoría sí fracasaría si los países que empezaron el desarrollo capitalista como países indígenas ahora fueran no indígenas; o si en ellos, manteniendo su condición de indígenas, la composición de los grupos sociales en el sector agrícola como campesinos, asalariados y capitalistas no estuviera asociada a la etnicidad. La presencia de grupos étnicos en la estructura social se distribuiría aleatoriamente. En este caso, contrariamente a lo que la teoría sigma predice, las condiciones iniciales no contarían.

En las recientes décadas, una significativa emigración de la población rural ha tenido lugar en América Latina. Generalmente, este fenómeno se explica por la gran diferencia del ingreso promedio entre las zonas urbanas y las rurales. Si bien este proceso de emigración rural no tiene nada de sorpresivo, lo que sí es inesperado es que la gente siga viviendo en las zonas rurales. ¿Por qué continúan allí si las diferencias con el ingreso de las zonas urbanas persisten? Una posible explicación proviene de la teoría de la exclusión: la mayor parte de la población indígena no es parte de la oferta laboral en los mercados urbanos de trabajo.

Bajo otras modalidades de producción en el pasado, las poblaciones indígenas suministraban la fuerza de trabajo y, para los terratenientes, resultaba provechoso emplearlas. Pero ese sistema de producción (el sistema de la hacienda) no funcionaba con mercados laborales. Además, con el tiempo, la tecnología cambió y la producción capitalista se expandió, requiriendo particulares habilidades, incluyendo el dominio del castellano y una nueva ética del trabajo. Los trabajadores aptos para la producción capitalista, los trabajadores-y, aumentaron con el tiempo. Pero, como ya se mostró más arriba, las poblaciones indígenas fueron dejadas atrás.

El sistema de mercado está menos desarrollado en las zonas donde predominan las poblaciones indígenas

La mayor parte de los estudios empíricos han encontrado que las comunidades campesinas de América Latina desarrollan el intercambio monetario con el mundo externo. Económicamente no son comunidades aisladas. Sin embargo, dicha evidencia empírica no puede ser tomada como un indicador de que ellas se encuentran completamente integradas en el sector capitalista. Un estudio detallado sobre el intercambio que hacen las familias en una muestra de comunidades campesinas indígenas en los Andes peruanos, realizado por Figueroa (1984), mostró cuatro hallazgos:

(a) El intercambio económico de los bienes y el trabajo era realizado a través del intercambio dentro y fuera del mercado.
(b) El intercambio económico con otras comunidades indígenas y con los productores auto-empleados en pueblos provinciales era muy significativo.
(c) En tanto oferta de trabajo, los campesinos intercambian parte de su fuerza laboral con otros campesinos sobre bases de reciprocidad; otra parte es vendida a las fincas locales; y otra parte es vendida como trabajo temporal en los mercados formales de trabajo.
(d) En términos cuantitativos, este estudio calculó que, en promedio, las familias campesinas intercambian la mitad de su producto total y que el ingreso del trabajo representa un 25% del ingreso total. Parte de este intercambio es de mercado y parte de no-mercado.

El bien conocido argumento desarrollado por Schumpeter (1934) de que «el campesino vende su ternero tan astuta y egoístamente como el inversionista en la bolsa vende su portafolio de acciones» (p. 80) no parece ajustarse con lo que uno observa en las comunidades rurales indígenas de América Latina. Ese tipo de conducta se refiere a campesinos que venden sus pro-

ductos en un mercado urbano a unos compradores que ellos escasamente conocen. Pero esta no es su usual forma de intercambio. Dentro de su comunidad y dentro de su red social, los campesinos rara vez se comportan como inversionistas de bolsa. Allí predomina el intercambio que es de no-mercado.

El alto grado de integración entre los auto-empleados rurales y urbanos (entre los dos subsectores dentro del sector de subsistencia) parece ser un perfecto calce entre oferta y demanda. Dadas sus dotaciones de bienes económicos, incluyendo el capital humano, estos grupos solo pueden producir bienes de baja calidad; y dado sus bajos ingresos, estos son los bienes y servicios que gozan de la más alta demanda. Esto es consistente con la observación de que los sectores de subsistencia producen, más que nada, «bienes inferiores».

Si estos subsectores producen bienes inferiores y son pobres, son economías complementarias y pueden ser consideradas como una sola economía. Esto es precisamente el supuesto que se ha adoptado en este modelo de la economía sigma. El intercambio de bienes con el sector capitalista ha sido ignorado, pues se supone que este intercambio no es el factor esencial para entender el tamaño de este sector.

Debido a que el campesinado participa parcialmente en el intercambio de mercado, su ingreso real también tendría que depender de los precios relativos del mercado. Pero la teoría ignora el efecto de estos precios relativos sobre el proceso de crecimiento económico de la economía campesina. Aunque sí hay intercambio de mercado en la economía campesina, la teoría hace abstracción de este intercambio, como si no lo hubiera. La teoría asume que los cambios en los precios relativos pueden tener un efecto de nivel sobre el ingreso (un salto), pero no un efecto de crecimiento continuo en la productividad y el ingreso.[2]

[2] Lucas (1988) hace un supuesto similar en su teoría del desarrollo. Por ejemplo, él arguye que la liberalización del comercio —un cambio en los precios relativos originado por una reducción de las tarifas— tendría un efecto de nivel *(level effect)* y no un efecto de crecimiento *(growth effect)* en el ingreso nacional.

La evidencia empírica es clara en señalar que en Latinoamérica el sistema del mercado está menos desarrollado en las zonas rurales donde las poblaciones indígenas son predominantes. Para el campesinado, los costos de transacción con el mercado parecen ser más altos que aquellos mismos dentro de sus redes sociales. Los bienes públicos (sistemas de transporte y comunicaciones) son necesarios para el desarrollo del mercado. No obstante, los Gobiernos invierten muy poco en la provisión de esos bienes públicos en zonas donde la población indígena es predominante. Esto es consistente con la hipótesis de la exclusión política. Y sin mercados desarrollados será muy caro para el campesinado lograr crecimiento económico.

Las innovaciones tecnológicas son limitadas en las zonas donde las poblaciones indígenas son predominantes

Estudios empíricos sobre el campesinado latinoamericano han encontrado que hay limitaciones de oferta y demanda en lo que concierne a la adopción de innovaciones tecnológicas. Solo una pequeña fracción del campesinado ha adoptado nuevas tecnologías.

Un estudio llevado a cabo dentro del Programa ECIEL[3] investigó trece microrregiones de comunidades campesinas, en cuatro países (Brasil, México, Paraguay y Perú) con el fin de explicar esa baja tasa de adopción. La conclusión fue que dichas adopciones requieren altos niveles de capital de trabajo y capital humano, de los que carecen los campesinos. La nueva tecnología es más intensiva en capital de trabajo y en capital humano. El incremento de la productividad a través de la utilización de tecnología moderna requiere capacidad de cálculo numérico, alfabetismo y dominio del castellano. Pero toma de

[3] ECIEL son las siglas de Estudios Conjuntos para la Integración Económica Latinoamericana. Esta fue la red de investigación más importante de la región durante los setenta y ochenta.

6 a 7 años (más que toda la escolaridad primaria) para lograr tales calificaciones. La adopción de innovaciones, entonces, requiere niveles de educación que la mayoría del campesinado no posee. Los campesinos con educación superior a la primaria —mayormente ofrecida en las ciudades— son los adoptadores de la nueva tecnología (Figueroa 1986).

La segregación de las comunidades campesinas rurales hace del aprendizaje en la escuela una experiencia muy costosa en términos de recursos y tiempo. No solo la escuela rural es de baja calidad, sino que los ambientes rurales hacen muy difícil el logro de las destrezas numéricas y el dominio del castellano tan necesarios para participar en el proceso de modernización tecnológica. Como resultado, el cambio tecnológico en las comunidades campesinas procede a paso muy lento. Los jóvenes tratan de escapar a la segregación trasladándose a la ciudad, acumulando capital humano allí. El hecho es que la cantidad y calidad de las escuelas públicas son muy bajas en las zonas donde predominan las poblaciones indígenas. El bajo nivel de escolaridad que se observa en esas zonas es consistente con la hipótesis de exclusión política y cultural. En consecuencia, la teoría sigma explica por qué es tan difícil reducir el núcleo duro de la pobreza rural en América Latina.

El hecho es que no hay convergencia en la productividad media entre las unidades capitalistas y las unidades campesinas en la mayoría de los países latinoamericanos. A la luz de la teoría sigma, la carencia de capital humano en la economía campesina es la causa próxima de este fenómeno. ¿Por qué existe tal carencia? Los campesinos constituyen la población-z en América Latina. La exclusión social es la causa última. El potencial para la modernización agrícola se encuentra bloqueada por la exclusión social.

Los precios relativos de mercado, incluyendo el salario real, no son los factores esenciales en el atraso de la economía campesina, como usualmente se señala. Aun cuando la economía campesina se encuentra parcialmente integrada al sistema de mercado, este intercambio ha sido ignorado en la teoría sigma,

porque se considera que los precios relativos tienden a oscilar entre subidas y bajadas en el largo plazo. El factor esencial se encuentra en el desarrollo del sistema de mercado, más que en los cambios en los precios relativos.

En suma, la evidencia que sobre Latinoamérica ha sido presentada aquí parece ser consistente con las predicciones de la teoría sigma. Los países que son étnicamente más heterogéneos también son más desiguales. Dentro de los países, las zonas en que predominan las poblaciones indígenas muestran una significativa economía campesina, un sistema de mercado poco desarrollado y tasas muy bajas de modernización tecnológica. En todos los países, la población indígena se ubica entre los más pobres. Estos resultados empíricos parecen sugerir que la población indígena se ajusta a la noción de trabajadores-z. Las predicciones empíricas derivadas de este modelo han probado también ser consistentes con la dualidad tecnológica observada en la agricultura latinoamericana. No hay razón para rechazar la teoría sigma por el momento.

La persistencia de la diversidad étnica muestra que la metáfora del «crisol» no se aplica a los países latinoamericanos. Pero la diversidad étnica no es un problema *per se;* es la jerarquía de esta diversidad lo que constituye un problema para el desarrollo económico. Esta jerarquía se ha convertido en parte de los mecanismos de segregación o exclusión social. Aquellos con activos culturales subvalorados por la cultura dominante fueron excluidos ayer, como resultado de lo cual tienen limitaciones para acumular capital humano y, por esa razón, todavía son excluidos hoy, y continuarán siendo excluidos mañana, en la medida en que se mantenga la cultura jerárquica.[4]

---

[4] Durante mi estadía como Profesor Visitante en la Universidad de Texas en Austin, en el otoño de 1997, asistí a un acalorado debate nacional sobre el racismo en EE.UU., iniciado por un profesor de leyes de esta universidad. Él dijo lo siguiente: «Los americanos negros y los méxico-americanos no pueden com-

6.7. Conclusiones

El modelo estático de la teoría sigma predice que la exclusión social, una vez instalado en el sistema, no puede ser eliminada endógenamente. La sociedad mostrará un equilibrio donde la heterogeneidad social es jerarquizada y así se reproduce periodo tras periodo. La exclusión social es el resultado de un evento histórico, de un choque fundacional, como es el caso de la dominación colonial. Esta condición inicial tiende a reproducirse.

En la sociedad sigma, existen barreras a la movilidad social de los pobres. La desigual dotación de activos sociales hace que existan marcadores sociales que conducen al bloqueo en el progreso de los grupos de bajos ingresos; dichos marcadores se basan en la etnicidad. Los pobres tienen dificultades para salir de su situación porque nacieron en la familia equivocada, con sexo equivocado o en el lugar equivocado.

Hasta aquí se han estudiado tres sociedades abstractas —épsilon, omega y sigma— bajo modelos teóricos estáticos. La inversión privada ha sido considerada una variable exógena. Este supuesto implica que también el crecimiento del ingreso nacional está exógenamente determinado. Las predicciones de estos modelos son consistentes con los datos básicos de las rea-

---

petir con los americanos blancos en términos académicos, debido a que sus culturas no condenan el fracaso». Yo estaba sorprendido de que nadie intentara explicar por qué «sus culturas» eran así. La teoría de la exclusión social diría que sus culturas son endógenas; es decir, un resultado del proceso de segregación. Por esos días leí un libro sobre multiculturalismo en América, de Glazer (1997), el que, implícitamente, da un gran respaldo a este punto de vista; un libro que también me mostró que la segregación de los negros en los EE.UU. tiene muchas similitudes con el caso de las poblaciones indígenas en América Latina. Glazer utiliza al matrimonio entre razas como un índice de segregación y encuentra que el 97% de las mujeres negras se casan con hombres negros. El hecho de que el coeficiente Gini de EE.UU. esté por sobre el promedio del primer mundo es también consistente con la teoría de la exclusión.

lidades que intentan explicar, es decir, el primer mundo se parece a la economía épsilon; el tercer mundo que no tienen un pasado de dominación colonial se parece a la economía omega; y el tercer mundo con una herencia colonial importante se parece a la economía sigma. Estos modelos también predicen viscosidad en los cambios en el grado de desigualdad en la distribución del ingreso nacional en cada una de las sociedades. Esta predicción es consistente con lo que muestran los estudios empíricos (Li, Squire y Zou 1998).

En el corto plazo, las teorías predicen que las variables nominales afectan las variantes reales, es decir, que el dinero es no-neutral. Esta predicción es consistente con la evidencia empírica. Un reciente estudio empírico basado en una muestra de 20 países del tercer y primer mundo encontró precisamente esta relación (Rand and Tarp 2002).

Otra predicción empírica que se deriva de estos tres modelos es que, con la acumulación de capital, las sociedades omega se transforman en sociedades épsilon, pero las sociedades sigma no pueden transformarse en sociedades épsilon. Esta predicción es también consistente con la experiencia histórica del desarrollo del capitalismo mundial (Maddison 1995). A esta etapa de este estudio no hay razón para no aceptar estas tres teorías.

La cuestión pendiente es la de explicar la falta de convergencia en cuanto al nivel de ingreso que se observa entre los países del primer mundo y del tercer mundo. Esta pregunta requiere construir modelos dinámicos de las tres teorías —modelos en donde la inversión privada sea endógena—. ¿Predicen estos modelos dinámicos una tendencia natural a la convergencia? Los capítulos siguientes se ocupan de este tema.

Capítulo 7
Modelos dinámicos

En los modelos estáticos de las economías épsilon, omega y sigma estudiados hasta ahora, el crecimiento del producto total en cada economía era exógeno, pues la inversión privada era exógena. Según esos modelos, es el «espíritu animal» de los capitalistas (como diría Keynes) el que genera el crecimiento económico. Si la inversión es exógena, el crecimiento de las economías sería apenas un evento. Los estudios empíricos muestran, en cambio, que existen ciertas regularidades en el crecimiento económico de los países (cf. Barro y Sala-i-Martin 1995). Por lo tanto, se necesita construir modelos dinámicos, en los que la inversión sea endógena.

Para transformar un modelo estático en otro dinámico se tiene que crear relaciones intertemporales entre algunas variables endógenas del sistema. Pero entonces se tiene que hacer un supuesto sobre la naturaleza del tiempo. En este estudio, se supondrá que el tiempo es discreto. Bajo este supuesto será mucho más simple mostrar que, en los tres modelos desarrollados en este estudio, el equilibrio dinámico no es sino una secuencia de situaciones de equilibrio estático.

La teoría neoclásica no podría explicar el crecimiento de los países debido a que en sus modelos estáticos ha sido refutada. Bastaría con este argumento para no insistir. Sin embargo, los modelos dinámicos de la teoría neoclásica se han desarrollado bastante y constituyen también el paradigma de la teoría del crecimiento económico. Una de sus principales predicciones es, sin embargo, inconsistente con los datos de la realidad. Dado que los países del tercer mundo están dotados con menos capital que los del primer mundo, la tasa de retorno del capital

sería mayor en el tercer mundo; luego, el capital debería moverse masivamente a los países del tercer mundo. Pero esto no es lo que se observa en la realidad.

El profesor Lucas (1990) ha tratado de rescatar la teoría introduciendo el capital humano en el modelo. Si se introduce el supuesto de que los países del primer mundo están más dotados no solo de capital físico sino, también, de capital humano, entonces la relación entre las tasas de retorno del capital entre los países del primer y tercer mundos no tiene que ser como la mencionada arriba, ni debe esperarse que el capital se mueva masivamente al tercer mundo. El modelo ahora establece que el capital físico se mueve entre las economías hasta que las tasas de retorno tiendan a igualarse en el margen.

El modelo de Lucas tiene dos problemas. En primer lugar, el modelo no ha sido puesto a falsación. En segundo lugar, debido a que el capital humano es resultado de un proceso de acumulación, la inversión en capital humano es también endógena y tendría que ser explicada. En realidad, como reconoce Jones (1998), la teoría neoclásica no ha producido todavía una teoría aceptada *(canonical theory)* de la inversión. Se hace necesario entonces desarrollar una teoría de la inversión privada.

7.1. Un modelo dinámico para la economía épsilon

El modelo estático de la economía épsilon presentado antes puede ser transformado en un modelo dinámico introduciendo varios supuestos auxiliares que sean consistentes. En términos del contexto, la economía épsilon está sujeta ahora a la incertidumbre. Los individuos enfrentan el problema de que las decisiones tomadas en el período uno generan resultados inciertos en el período dos.

Sobre la racionalidad de los capitalistas, se introducen dos supuestos. El primero de ellos consiste en que los objetivos perseguidos por los capitalistas son dos y están ordenados jerárquicamente.

El objetivo de mantener una posición social tiene un lugar más alto en la jerarquía que el objetivo de la maximización de beneficios. Para asegurar su posición social, los capitalistas buscan tener un nivel de riqueza por encima de un umbral que los convierte en capitalistas. Llamemos a este umbral K_b^*. Esta es la mínima cantidad de capital que un individuo debe poseer para seguir siendo capitalista (considérese la riqueza necesaria para ser elegible para un crédito bancario, tal como se presentó en el capítulo 4).

El capitalista no participa de juegos económicos en los que existe el riesgo de tener pérdidas de tal magnitud que reduzcan su riqueza más allá del nivel de dicho umbral, independientemente de las posibles ganancias resultantes. Si dicha pérdida ocurriese, le significaría el desastre económico: el individuo dejaría de ser capitalista. Existen límites al riesgo que puede ser tomado por los capitalistas. Por ello, los capitalistas solo participan de aquellos juegos en los que los resultados excluyen posibles desastres económicos.

Los capitalistas que incurren en pérdidas en un período determinado sufrirán una reducción en su *stock* de capital. Pero estas pérdidas tienen un límite: no deben arriesgar el nivel K_b^*. Luego, dada su dotación inicial y el valor de K_b^*, existe un máximo flujo de pérdidas que el capitalista puede absorber. Para el capitalista *j*, llamemos a este valor (P_j^*). El riesgo de sufrir pérdidas u obtener ganancias viene de la actividad de producir bienes. El capitalista no elegirá opciones de producción que le signifiquen el riesgo de perder P_j^*. Estas opciones no pertenecerán a su conjunto factible de beneficios y riesgos, aunque la posibilidad de beneficios fuese mayor que en las otras opciones.

Existe también un riesgo de destrucción del *stock* de capital en cada período. Este riesgo puede ser asegurado por medio del mercado de seguros. El capitalista adquirirá la cantidad $P_b K_b^*$ dólares de seguro y deberá pagar una prima de $s P_b K_b^*$ por unidad de tiempo, donde *s* es la tasa de seguro.

El objetivo de mantener su posición social implica que el capitalista debe imponer límites a las pérdidas en que pueda incurrir:

$$-P_j^* = (K_b^* - K_{bj}) - s\,K_b^* = G\,(K_{bj}), \qquad G'>0 \tag{7.1}$$

Dado K_b^* (un umbral uniforme para todos los capitalistas) y la tasa de seguro, el umbral de pérdidas tolerables del capitalista ($-P_j^*$) depende solamente de su *stock* de capital. A mayor *stock* de capital, mayores las pérdidas que el capitalista puede absorber.

El riesgo del flujo de ganancias puede ser entonces absorbido enteramente por el capitalista. Este riesgo no es un riesgo asegurable en el mercado de seguros, debido a un problema de incentivos, conocido como el problema de riesgo moral. Por lo tanto, el capitalista no participa de ningún *juego riesgoso,* es decir, de cualquier juego que lo pueda llevar a tener pérdidas más allá de ($-P^*$), hecho que le implicaría un desastre económico. Este tipo de juego es evitado, aun si fuesen posibles amplias ganancias en el caso de que el evento fuese favorable. El capitalista es «adverso a juegos riesgosos» en este sentido preciso.

Esta formulación teórica que puede ser llamada *teoría de aversión a los juegos riesgosos* es diferente de la teoría convencional de *aversión al riesgo.* En esta última, las decisiones tomadas bajo condiciones de incertidumbre dependen de las preferencias del individuo; es más, el individuo no impone límites a las pérdidas en que pueda incurrir. En la teoría que introduce aversión a juegos riesgosos, en contraste, las decisiones de portafolio se determinan de manera jerárquica: en primer lugar, por las restricciones del individuo —su capacidad económica de absorber pérdidas—, que determinan el conjunto factible de juegos que el individuo puede jugar; y solo entonces, las preferencias determinan la elección dentro de este conjunto factible de juegos (véase el apéndice para la presentación formal de esta teoría).

El segundo supuesto acerca del comportamiento de los capitalistas dice que el flujo de beneficios será añadido enteramente a la acumulación de capital. Los capitalistas no consumen. Se supone que la proporción de los beneficios que dedican al

consumo es tan pequeña que puede ser ignorada; es *como si* no consumieran.

Estos dos supuestos sobre la racionalidad económica del capitalista pueden ser escritos de la siguiente manera. El capitalista *j* busca

$$\text{maximizar: } P_j' = P_b Q_{bj} - P_h D_{hj} - P_c D_{cj} - r D_{rj} - s P_b K_b^*,$$

sujeto a:

$$\begin{aligned} Q_{bj} &= f_j (D_{hj}, K_{bj}) \\ &= D_{cj} / n \\ &= D_{rj} / q P_b \\ K_{bj} (t+1) &= K_{bj} (t) + P_j (t) - FI_j \\ K_{bj} &> K_b^* \end{aligned} \qquad (7.2)$$

Los elementos en la función de producción son todos contemporáneos y, por esa razón, no están fechados. La inversión de hoy se suma al *stock* de capital de hoy para obtener el *stock* de capital de mañana. Los beneficios son, en parte, destinados a la inversión neta (I) y, por lo tanto, a aumentar el *stock* de capital de la economía épsilon del siguiente período. En una economía internacional abierta, con perfecta movilidad de capitales, los beneficios pueden ser también invertidos en el extranjero. Esta cantidad FI_j es, por lo tanto, deducida de los beneficios en la determinación del *stock* de capital de la firma para el siguiente período.

En este capítulo se supondrá que no existe movilidad del capital entre sociedades y, por lo tanto, toda la ganancia de la economía se invierte localmente. En el capítulo siguiente se tratará el caso de crecimiento con movilidad de capitales.

Al agregar el comportamiento de los capitalistas, se obtiene las siguientes relaciones:

$$\begin{aligned} Q_b (t) &= F (D_h (t), K_b (t)) \\ I (t) &= P (t) \\ K_b (t+1) &= K_b (t) + I (t) \\ Q_b (t+1) &= F (D_h (t+1), K_b (t+1)) \end{aligned} \qquad (7.3)$$

La primera ecuación es solo la agregación de la función de producción de las firmas individuales, donde Q_b es el producto neto (neto de depreciación y de exportaciones requeridas para importar D_c), y los costos de crédito y de seguro no son tomados en cuenta. La segunda ecuación dice que la inversión proviene de los beneficios domésticos. La tercera ecuación es solo la identidad de acumulación de capital. La cuarta ecuación es la función de producción con valores para el siguiente período.

Considérese nuevamente el modelo del equilibrio general estático que se estableció para la economía épsilon en el capítulo 4. Como se vio allí, la solución de equilibrio general estático de hoy determina el producto y los beneficios (entre otras variables endógenas) de hoy. Ahora se añade el supuesto de que los beneficios de hoy determinan la inversión de hoy, que determina el *stock* de capital de mañana y la solución de equilibrio general de mañana. Esta, a su vez, determina el producto y los beneficios de mañana, y así sucesivamente. Luego, los beneficios de hoy determinan el producto de mañana.

Dado que se está introduciendo relaciones intertemporales en las ecuaciones estructurales del sistema estático de equilibrio general y que este sistema tiene un equilibrio que es único y estable en cada periodo, entonces hemos transformado un sistema estático en uno dinámico. El *stock* de capital aparece ahora como condición inicial. El equilibrio dinámico de toda variable endógena es, entonces, una secuencia de situaciones de equilibrio estático.

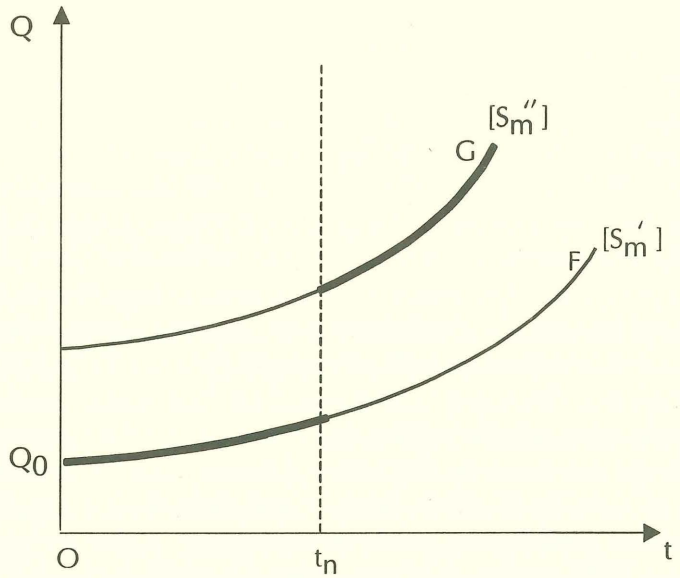
Bajo esta transformación, la forma reducida del sistema estático para cada variable endógena puede ser usada para definir su propio equilibrio dinámico. Considérese la forma reducida del modelo estático y abierto para la economía épsilon presentado en el capítulo 4 para el producto total —ecuación (4.18)—. Luego, el equilibrio dinámico del producto total puede ser escrito como sigue:

$$Q_b^{\circ}(t) = F\ (K_{b0}, \delta_0, S_h, S_m, z^*, r^*; t) \qquad (7.4)$$

Esta ecuación se refiere a valores del equilibrio del producto total en cada periodo de tiempo, manteniendo fijas las variables exógenas. Este es el equilibrio dinámico.

El equilibrio dinámico del producto total se representa por la curva F en la Figura 7.1. Aquí, el producto crece en el tiempo, con una trayectoria particular en el tiempo, en la que las variables exógenas permanecen constantes. Esta trayectoria es una curva creciente que no converge a un valor estacionario.

Figura 7.1
Efecto de la oferta de dinero sobre la trayectoria del producto en el tiempo



El equilibrio dinámico de cada variable endógena del sistema que no sea la desigualdad en el ingreso seguirá, en consecuencia, una trayectoria particular a través del tiempo para valores dados de las condiciones iniciales y de las variables exógenas. En este sistema dinámico, por lo tanto, el conjunto de variables puede ser clasificado como sigue:

Variables endógenas : $Q_b(t), D_h, K_b(t), I(t), w(t), W(t), P(t)$
Variables de condición inicial: K_{b0}, δ_0
Variables exógenas : S_h, S_m, z^*, r^*, t

Ciertamente, cambios en las variables exógenas alterarán las sendas temporales de las variables endógenas. Dado que hay estabilidad en el equilibrio para cada período (hecho que proviene de la propiedad del equilibrio en el sistema estático correspondiente), el método de la dinámica comparativa puede ser aplicado para derivar proposiciones beta.

¿Cuál es el efecto de un cambio en una variable exógena en la trayectoria del producto? Un cambio en una variable exógena afectará la solución estática de equilibrio general del período, incluyendo los beneficios totales. Con esta nueva cantidad de beneficios, la acumulación de capital y el producto seguirán una nueva trayectoria. Esta variable exógena afecta el producto no solo en el período de su cambio sino, también, en su trayectoria. En la ecuación (7.4), la dinámica comparativa indica que la curva que muestra el equilibrio dinámico se desplazará hacia arriba cuando aumenta cualquiera de las variables exógenas, a excepción de la tasa de interés internacional (efecto negativo) y el aumento en la oferta laboral (efecto neutro).

El crecimiento económico de largo plazo no es independiente de los *shocks* de corto plazo de la economía, tales como cambios en la oferta de dinero o en los términos de intercambio. Por ejemplo, un aumento en la oferta de dinero eleva los beneficios en el mismo período. Por lo tanto, se acumulará más capital para el siguiente período, y el producto será mayor. Habrá un desplazamiento hacia arriba en la trayectoria del producto total. Claramente, la economía épsilon tendrá, en cada período, mayor producto en relación con la trayectoria inicial. El efecto sobre la pendiente de la nueva trayectoria, que indica una mayor o menor tasa de crecimiento, no puede ser derivado del modelo.

Esta proposición beta se muestra gráficamente en la Figura 7.1. El equilibrio dinámico inicial está dado por la curva F. En el

período t_n, la cantidad de dinero aumenta y el nuevo equilibrio dinámico es dado por la curva G. Este modelo teórico no predice necesariamente cambios en las tasas de crecimiento del producto. Tampoco predice un valor de estado estacionario; es decir, cada trayectoria tiene un punto de partida pero no un punto de llegada. Un equilibrio dinámico con un valor de estado estacionario es una propiedad que la mayoría de teorías de crecimiento económico persiguen. Incluso buscan que el punto de llegada sea independiente del punto de partida.[1]

¿Cuál será la trayectoria de la distribución del ingreso? En el proceso de acumulación de capital, más cantidades de capital físico son incorporadas al proceso de producción. La productividad del trabajo crece, como también lo hace la curva de demanda de trabajo. En el modelo estático desarrollado anteriormente, el efecto de la acumulación de capital sobre la ratio de beneficios a la masa salarial era ambiguo. El equilibrio dinámico, una secuencia de equilibrios estáticos en cada período, describe también una trayectoria ambigua.

Si la oferta monetaria aumenta en un período, el efecto sobre la ratio P/W en ese mismo periodo será ambiguo (esto viene de la estática comparativa presentada en el capítulo 4). Luego, la nueva trayectoria de P/W en el tiempo será también ambigua. La misma conclusión puede derivarse de cambios en los términos de intercambio internacionales o en la tasa de interés internacional. Debido a que son todos ambiguos en el equilibrio estático, sus efectos sobre la ratio P/W no pueden traducirse en un nuevo equilibrio dinámico particular, es decir, en una nueva

[1] Los modelos de crecimiento neoclásico hacen el supuesto de que la ratio de ahorros es constante; aun más, esta ratio es aplicada al producto bruto, no al producto neto (neto de depreciación). Luego, en ausencia de cambio tecnológico, la trayectoria de este sistema dinámico muestra valores de estado estacionario. El problema de la convergencia de niveles de ingreso apareció por primera vez en este contexto. Si la ratio de ahorros se asume fija pero respecto al producto neto (que la sociedad es libre de consumir o ahorrar), no se obtendría ese valor de estado estacionario.

trayectoria particular. Por lo tanto, *shocks* de corto plazo (internacionales o domésticos) no tienen un efecto definitivo en el equilibrio dinámico de la desigualdad del ingreso.

En suma, el modelo dinámico de la economía épsilon predice una trayectoria particular a través del tiempo para el producto total, que se desplazaría hacia arriba al aumentar la oferta de dinero y los términos de intercambio, y se desplazaría hacia abajo al aumentar la tasa de interés internacional y la oferta laboral. Pero este modelo no predice una trayectoria particular para el caso de la desigualdad en el ingreso; además, los cambios en las variable exógenas no modifican esta trayectoria. El equilibrio dinámico de la desigualdad en el ingreso puede ser representado como una curva horizontal a través del tiempo, la misma que permanecería inalterada ante cambios en las variables exógenas.

Este modelo dinámico, sin embargo, no puede predecir una trayectoria de beneficios que sea creciente en el tiempo. Los datos empíricos muestran, en contraste, que los beneficios se elevan en el tiempo y que la participación de beneficios en el ingreso nacional es invariable en cada país. Este es uno de los hechos estilizados famosos de Nicholas Kaldor acerca del proceso de crecimiento económico en el primer mundo. Barro y Sala-i-Martin (1995) han demostrado que la mayoría de estudios empíricos muestran conformidad con las proposiciones de Kaldor (Barro y Sala-i-Martin 1995: 5). El cambio tecnológico ha sido ignorado en este modelo, y dicho supuesto será relajado ahora.

7.2. Progreso tecnológico

La aproximación usual al análisis del progreso tecnológico es considerar que las innovaciones pueden tomar tres formas. Estas innovaciones pueden permitir a las firmas producir la misma cantidad de producto con menos cantidad relativa de capital (inno-

vaciones ahorradoras de capital) o con menor cantidad relativa de trabajo (innovaciones ahorradoras de trabajo) o con menor cantidad proporcional de ambos (innovaciones neutrales). Estas innovaciones son incorporadas en la función de producción a manera de nuevas funciones de producción y, gráficamente, a modo de desplazamientos en la curva de producción. Sin embargo, este enfoque no ayuda a establecer proposiciones beta.

La función de producción es una proposición no observable; pertenece a los supuestos tipo alfa. Empíricamente, uno puede observar solamente un método de producción, el que está en uso, mas no todo el conjunto de posibles métodos (este fue el problema con las tablas de insumo-producto que construyó Leontief, en las que toda la tecnología de la economía se reducía a un único vector: el observado). La tecnología es un *elemento exógeno,* pero no es *una variable exógena* en el proceso económico. Por lo tanto, cambios en la tecnología y en la función de producción tampoco son observables. Decir que cierta variación en las variables reales se debe a un cambio tecnológico es tautológico, debido a que dicho efecto puede siempre ser «explicado» invocando el tipo apropiado de cambio tecnológico que pudiese haber ocurrido (ahorrador de trabajo, ahorrador de capital o neutral). Esta proposición no es falsable.

¿Cómo puede transformarse la tecnología en una variable exógena en nuestro modelo de una economía de un solo bien? En esta economía abstracta, el bien B es fabricado a partir del bien B (acero a partir de acero, cereales a partir de cereales). Un procedimiento simple es considerar a las innovaciones tecnológicas como nuevos bienes de capital, fabricados a partir del mismo bien B, pero que son cualitativamente diferentes, más productivos (acero en forma de un robot, en vez de un tractor o de un arado; variedades de cereales de alto rendimiento, en vez de cereales nativos). Este nuevo bien de capital incorpora innovaciones derivadas de nuevo conocimiento científico y tecnológico. Esta característica de los nuevos bienes de capital es observable. El efecto de este cambio en la variable exógena —el

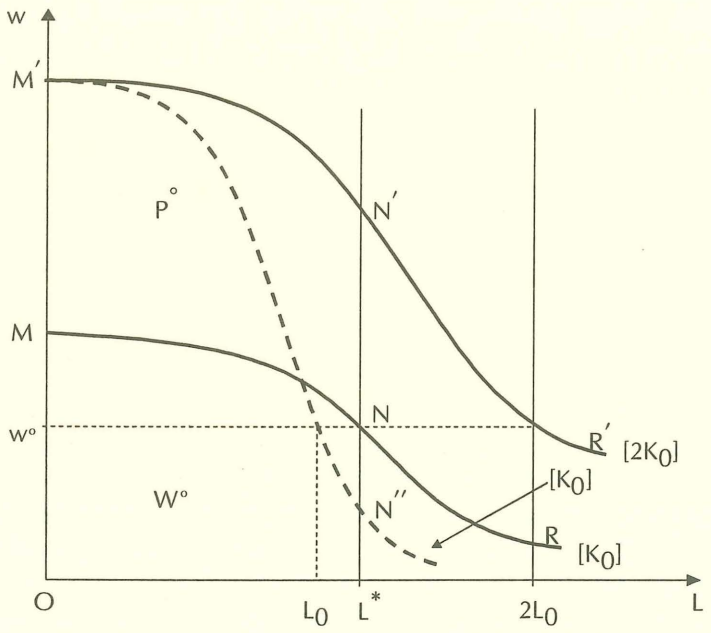
surgimiento de un nuevo bien de capital, como las computadoras— sobre las variables endógenas es factible de determinarse por medio del análisis estadístico.

Las firmas pueden usar sus beneficios para invertir no solo en capital físico sino, también, en investigación para producir nuevas innovaciones tecnológicas. De manera consecuente con su racionalidad maximizadora de beneficios, ¿qué tipo de innovaciones buscarían obtener los capitalistas?

Para un nivel de empleo dado (digamos $D_h = L^*$), es fácil reconocer que una innovación ahorradora de trabajo genera mayores posibilidades de beneficios en relación con otras formas, como se puede ver con la ayuda de la Figura 7.2. Con menos trabajo, la firma puede producir la misma cantidad de producción porque el nuevo capital sustituye la energía humana o el poder del cerebro humano. Por lo tanto, la productividad promedio y la productividad marginal del trabajo aumentan. La curva de demanda de trabajo se desplaza hacia fuera, aunque no paralelamente, con el mayor aumento a menores niveles de empleo.

Para una innovación ahorradora de trabajo significativa, la nueva curva de demanda puede cruzar a la anterior desde arriba. En este caso, al salario inicial, la cantidad de trabajo demandada se reduce. Esta situación es también representada en la Figura 7.2. Allí, MN es la curva de demanda de trabajo inicial, dado el *stock* de capital. Supóngase que la innovación ahorradora de trabajo es tal que la nueva curva de demanda de trabajo M'N'' (donde el *stock* de capital permanece constante) corta a la anterior por arriba. Al nivel inicial de empleo, la productividad marginal del trabajo cae de N a N''. Luego, la cantidad de empleo tendría que disminuir; y a este nuevo nivel de empleo, los beneficios aumentan y la masa salarial cae.

Figura 7.2
Efecto de la inversión sobre la curva de demanda de trabajo



En el caso de una innovación ahorradora de capital, la curva de demanda también se desplaza hacia fuera, pero la nueva curva se abre hacia afuera. Un nuevo tipo de capital humano reemplaza al anterior. La productividad marginal del trabajador aumenta. Por lo tanto, al salario inicial, el nivel de empleo aumenta, los beneficios caen y la masa salarial aumenta.

De manera consistente con su racionalidad maximizadora de beneficios, los capitalistas prefieren innovaciones ahorradoras de trabajo. Si los capitalistas invierten en la generación de nuevas tecnologías, buscarán innovaciones ahorradoras de trabajo. Los nuevos bienes de capital serán ahorradores de trabajo. Esta es la predicción básica de la naturaleza de las innovaciones tecnológicas. La otra predicción es que, debido a las eco-

nomías de escala existentes en la actividad de búsqueda de nuevas tecnologías, solo las firmas más grandes invertirán en la generación de tal tecnología (inversión en investigación y desarrollo, conocido como «R&D» por sus siglas en inglés). Por lo tanto, las innovaciones tecnológicas son endógenas a las firmas grandes, pero exógenas al resto.[2]

En suma, la teoría de maximización de beneficios predice que el progreso tecnológico tomará la forma de innovaciones ahorradoras de trabajo. Sobre la evidencia empírica respecto a que las tecnologías son ahorradoras de trabajo, Wassily Leontief escribió una vez diciendo que la primera revolución industrial se basó en nuevos bienes de capital que básicamente sustituyeron a la energía humana y animal, pero que los mayores perdedores fueron los animales. Como consecuencia, el *stock* de caballos en Europa declinó significativamente. La revolución tecnológica presente está sustituyendo principalmente el poder del cerebro humano. Leontief trajo a colación la cuestión de cómo el sistema social se ajustará, dado que los humanos no son caballos. En general, el hecho estilizado parece ser que la mayoría de los nuevos bienes de capital sustituyen trabajo humano.

Si el cambio tecnológico es ahorrador de trabajo, los beneficios se elevarán en este proceso de progreso tecnológico. Este es el efecto puro de las innovaciones tecnológicas, manteniendo constante el nivel de capital. Sin embargo, los capitalistas invierten también en la acumulación de capital físico, donde las innovaciones tecnológicas están incorporadas en los nuevos bienes de capital.

[2] Se asumió en el capítulo 3 que las firmas no conocían la tecnología para producir nuevas tecnologías. Por lo tanto, no podían producir innovaciones ahorradoras de trabajo. En la realidad, las firmas pueden realizar inversiones en investigación y desarrollo para desarrollar nuevas tecnologías. Sin embargo, el producto de este proceso de investigación no es un flujo continuo de nuevos bienes de capital, como en el proceso de producción de zapatos. Este era el significado de ese supuesto. Ahora se supone que las firmas invierten en R&D y persiguen en el largo plazo, entre otros objetivos, innovaciones ahorradoras de trabajo.

La inversión, en consecuencia, traslada la curva de demanda de trabajo de dos maneras: una es el efecto de la acumulación de capital, manteniendo la tecnología constante; la otra es el efecto de la mejora de la calidad del nuevo bien de capital —mejora producto del progreso tecnológico incorporado en la nueva cantidad de capital—. Este supuesto implica que, aun en el caso de una innovación ahorradora de trabajo, la inversión desplazará la curva de demanda de trabajo hacia afuera. Sin embargo, este desplazamiento no será paralelo sino, más bien, sesgado hacia mayores desplazamientos a menores niveles de empleo y pequeños desplazamientos a altos niveles de empleo. En consecuencia, en el proceso de crecimiento económico endógeno, los beneficios crecerán.

Estos efectos se muestran en la Figura 7.2. Como se indicó más arriba, M'N'' representa el desplazamiento de la curva de demanda de trabajo, debido únicamente al progreso tecnológico, es decir, manteniendo constante el *stock* de capital (la misma cantidad de acero en un robot, que es más productivo que el tractor). Al incluir el aumento del *stock* de capital (más acero), la nueva curva de demanda de trabajo se convierte en M'N'. Este es un desplazamiento hacia fuera de la curva de demanda de trabajo, pero no es paralelo si la función de producción agregada muestra retornos constantes a escala (y este debe ser el caso en un mundo de un solo bien y competencia perfecta). En el nivel de empleo L*, este efecto combinado genera un aumento tanto de los beneficios como de los salarios. El cambio en la distribución del ingreso, la ratio P/W, queda, por tanto, indeterminado.

En el equilibrio general, la interacción entre la inversión y el crecimiento del producto a través de la acumulación de capital pura está incorporado en la ecuación (7.4). Esta interacción que incluye el cambio tecnológico puede ser introducida en el equilibrio dinámico como sigue:

$$Q_b^{\circ}(t) = A(t)\, F(K_{b0}, \delta_0, S_h, S_m, z^*, r^*; t)$$
$$A(t) = f(I(t)) \qquad (7.5)$$

La forma reducida resultante para la producción y la distribución es:

$$Q_b^{\circ}(t) = F'(K_{b0}, \delta_o, S_h, S_m, z^*, r^*; t) \qquad (7.6)$$
$$D^{\circ}(t) = G(\delta_0)$$

El progreso de la tecnología se representa por A(t). Pero este es endógeno, pues depende de la cantidad de inversión. Luego, la función F'captura ambos efectos, el de mayor capital y nueva tecnología incorporada en el nuevo bien de capital. En la forma reducida, esta función no se diferencia de la función F en (7.4) debido, precisamente, al supuesto de que el cambio tecnológico es endógeno.

El equilibrio dinámico de la desigualdad del ingreso es representado ahora por la función G. La trayectoria en el largo plazo tiende a ser constante como resultado de los efectos contrapuestos en los valores de equilibrio de los precios y cantidades. Luego, la desigualdad inicial en la distribución de activos determina el grado de desigualdad de los flujos de ingresos en los periodos siguientes.

En resumen, al introducir innovaciones ahorradoras de trabajo en el modelo dinámico, se puede derivar las siguientes proposiciones beta: los beneficios totales aumentan en el proceso de crecimiento económico y la distribución entre beneficios y salarios tiende a ser constante. Estas predicciones están ahora de acuerdo con los datos empíricos sobre nivel de beneficios y distribución del ingreso mencionados por Barro y Sala-i-Martin (1995). Dado que los beneficios crecen en el tiempo, la inversión aumenta en el tiempo y también lo hacen los niveles de ingreso. Dado el nivel de oferta de trabajo (o suponiendo que la población crece a una tasa menor que el producto total), el ingreso per cápita también crece en el tiempo.

7.3. Un modelo dinámico de la economía omega

La economía omega se compone de un sector capitalista y de un sector de subsistencia. Se supone que el sector capitalista opera como en la economía épsilon, excepto que el mecanismo de extracción del esfuerzo de los trabajadores se basa en el salario relativo y no en la tasa de desempleo. A pesar de esta diferencia en las reglas de la disciplina laboral, el resultado analítico es que la forma reducida del sistema estático en la economía épsilon es igual a la del sistema estático de la economía omega. En efecto, la matriz de proposiciones beta, que se obtienen de la forma reducida, que corresponde a la economía épsilon en el Cuadro 4.1, es exactamente igual a la que corresponde al sector capitalista de la economía omega en el Cuadro 5.1.

Otro supuesto sobre la economía omega es que la acumulación de capital tiene lugar solo en el sector capitalista. El sector de subsistencia es residual al equilibrio estático que se obtenga en el sector capitalista.

El modelo estático de omega se puede transformar en un modelo dinámico si se incluye el supuesto de que la inversión es endógena. Debido a que este sistema dinámico solamente se circunscribe al sector capitalista, este sería igual al sistema dinámico desarrollado en la economía épsilon. La solución de este modelo dinámico ya la conocemos y puede entonces ser aplicada al sector capitalista de la economía omega. El sector de subsistencia continuará siendo residual al equilibrio dinámico del sector capitalista.

El proceso dinámico en la economía omega puede entenderse con la ayuda de la Figura 5.1, panel (c). Con la inversión en capital físico, la curva de demanda de trabajo MR se desplazará hacia fuera a través del tiempo. En este continuo desplazamiento, llegará a colocarse totalmente por encima de la curva de productividad marginal del sector de subsistencia *mr* y cortarla por encima del punto *m*. La sobrepoblación desaparecería

y, con ella, también el sector de subsistencia. El proceso de acumulación de capital y progreso tecnológico generará una transformación social. Omega no solamente devendrá en una sociedad más rica, con salarios más elevados, sino que cualitativamente se transformará en una sociedad épsilon, una sociedad capitalista pura.

El equilibrio dinámico en la distribución del ingreso es una trayectoria horizontal, sin tendencia a la subida o a la bajada en el tiempo. Este resultado se puede entender con facilidad si se recuerda que el equilibrio dinámico es la secuencia de situaciones de equilibrio estático y que, en el modelo estático de la economía omega, la desigualdad no cambia ante aumentos en el *stock* de capital (capítulo 5).

Las ecuaciones que muestran el equilibrio dinámico tanto en el ingreso nacional (Y) como en el grado de desigualdad en su distribución (D) se pueden escribir de la siguiente manera:

$$Y^{o}(t) = F(K_{b0}, \delta_0, S_h, S_m, z^*, r^*; t) \qquad (7.7)$$
$$D^{o}(t) = G(\delta_0)$$

La dinámica comparativa puede ser derivada directamente de estas ecuaciones. En el caso del ingreso nacional, los efectos de todas las variables exógenas son positivos, incluyendo la variable tiempo *t,* con la sola excepción de la tasa de interés internacional y la oferta laboral cuyos efectos son negativos. En el caso de la desigualdad, el equilibrio dinámico es un valor constante, que solamente depende de la desigualdad inicial en la dotación individual de activos económicos.

La reproducción de la desigualdad periodo tras periodo en la economía omega puede parecer paradójica. El modelo utilizado supone que todos los individuos son socialmente homogéneos y están, además, dotados con iguales cantidades de capital humano. Los trabajadores auto-empleados podrían, en un contexto de desarrollo, adoptar los nuevos bienes de capital que se están produciendo en el sector capitalista. No tienen ninguna

restricción de capital humano para hacerlo. La curva de productividad del trabajo en el sector de subsistencia se elevaría periodo tras periodo de manera endógena. Los ingresos en este sector se elevarían y el salario real tendría que aumentar para mantener la disciplina laboral en las firmas capitalistas.

Este proceso no puede darse debido a la exclusión económica que enfrentan los trabajadores auto-empleados. No solo están excluidos del mercado laboral sino, también, de los mercados de crédito y de seguros. La reproducción de la desigualdad es resultado del mecanismo de la exclusión económica, que opera a través de estos tres mercados. Debido a este papel, estos mercados tienen una mayor jerarquía que el resto de los mercados y, por eso, se les ha denominado mercados básicos.

El progreso tecnológico en el sector capitalista no se desparrama sobre el sector de subsistencia. El efecto es más indirecto. Al crecer el sector capitalista aumenta el nivel de empleo y, por lo tanto, se reduce el exceso de oferta de trabajo. La menor cantidad de auto-empleados conduce a un aumento en el ingreso medio en el sector de subsistencia. Pero, en este proceso, la desigualdad en los ingresos difícilmente puede cambiar.

7.4. Un modelo dinámico de la economía sigma

El sistema dinámico en la economía sigma es mucho más fácil de entender, una vez que los modelos dinámicos en las economías épsilon y omega han sido resueltos. La economía sigma no es sino una combinación de las economías omega y un sector de subsistencia —para los trabajadores socialmente excluidos— que no tiene relaciones económicas con el resto de la economía.

Para transformar el modelo estático de la economía sigma en un modelo dinámico, todo lo que se necesita es construir un modelo dinámico para la sub-economía omega contenida en sigma. Esta construcción ya ha sido hecha en la sección ante-

rior. Los resultados pueden ser aplicados ahora también a la economía sigma. Debido a la dualidad económica, el proceso de crecimiento que tiene lugar en el sector capitalista no tiene efectos, ni directos ni indirectos, sobre el sector de subsistencia de los trabajadores-z.

El ingreso nacional en la economía sigma incluye el ingreso que se genera en el sector de subsistencia-z. Pero este ingreso sectorial es exógenamente determinado y, además, no tiene efectos sobre las variables endógenas de la sub-economía omega. Luego, este ingreso se puede considerar ya incluido en la ecuación que determina el equilibrio dinámico del ingreso nacional. La primera ecuación del sistema (7.7) se aplica también a la economía sigma.

En cuanto a la desigualdad en la distribución del ingreso nacional, el bajo ingreso relativo de los trabajadores-z se une a una desigualdad que ya existe en la sub-economía omega y contribuye a generar un mayor grado de desigualdad en la economía. En el equilibrio dinámico, este grado de desigualdad no se modifica. Debido a que el ingreso permanece invariable en el sector de subsistencia-z, los trabajadores-z son dejados atrás. Las brechas de ingresos se amplían entre los trabajadores-z y el resto de la economía. Este resultado tendería a aumentar la desigualdad global. Pero este resultado no es una necesidad. En este proceso se da un cambio en los pesos de los grupos de trabajadores, como es el caso de la disminución en el tamaño del sector de subsistencia-y (el grupo más pobre dentro de la subeconomía omega), mientras que el tamaño del grupo de asalariados aumenta. Por lo tanto, las curvas de Lorenz pueden cruzarse.

En cualquier caso, este modelo dinámico predice para la desigualdad en los ingresos trayectorias que toman la forma de una curva horizontal o una curva creciente (aunque no muy pronunciada). Esta trayectoria se determina esencialmente por la desigualdad inicial en la dotación individual de activos. Como en el caso de la economía omega, las variables exógenas no

tienen efectos sobre la trayectoria. Por lo tanto, la segunda ecuación del sistema (7.7) también se aplica a la economía sigma.

7.5. Equilibrio y reproducción social

La economía épsilon opera con desempleo y con desigualdad en el ingreso. El desempleo es el mecanismo que asegura la disciplina laboral. La existencia de desempleo es, por lo tanto, fundamental para el funcionamiento de la economía capitalista; es necesaria para la generación de beneficios máximos. No son los bajos salarios reales los que constituyen la base de los beneficios y la acumulación, como señalaba la teoría clásica; no son los salarios reales, como indica la teoría de los salarios de eficiencia; es el desempleo el que juega dicho papel. Más aún, la inversión no puede eliminar el desempleo. Luego, la afirmación del historiador económico Garraty (1978), «el desempleo es la enfermedad del capitalismo», se mantiene cierta.

De acuerdo con la teoría épsilon, una economía con pleno empleo es una ilusión; el sistema capitalista necesita desempleo para operar. Como en el juego de las «sillas musicales», en el que hay n jugadores y n-1 sillas, no importa lo que los trabajadores hagan, siempre habrá trabajadores sin empleo; es más, debido a este mecanismo, los trabajadores jugarán el juego de la disciplina laboral.

Pero, ¿por qué debe persistir el desempleo? ¿Hay algo que los trabajadores puedan hacer al respecto? Los trabajadores desempleados tienen la opción de convertirse en auto-empleados. ¿Qué les impediría hacer esto en una economía competitiva? Para formar una empresa, los trabajadores necesitan capital. Los desempleados podrían alquilar servicios de capital si tal mercado existiese. Pero dicho mercado no existe porque los capitalistas prefieren usar su capital para generar beneficios en lugar de rentas.

Los desempleados tendrían que invertir en la acumulación de capital físico. Pero esta opción requiere financiamiento y los trabajadores no tendrán ahorros para financiar la inversión. Este resultado es independiente de los salarios que reciban, pues ellos se ven obligados (por el «efecto demostración») a aumentar su consumo continuamente para tratar de alcanzar el nivel de vida de los capitalistas. Si los trabajadores son capaces de ahorrar, sus ahorros se destinarán a reproducir el capital humano en sus hijos, no a la acumulación de capital físico. Los trabajadores deben ahorrar para mantener el capital humano constante a través de las generaciones. Pero, visto así, ese ahorro es bruto solamente.

¿Podrían los trabajadores obtener financiamiento por medio del crédito? Si el mercado de crédito operara como un mercado walrasiano, ninguna persona que deseara obtener crédito al precio de mercado quedaría fuera. La demanda de crédito que viene de los desempleados estaría siempre en equilibrio. Los trabajadores poseerían capital y la tasa de desempleo sería nula. Pero el mercado de créditos es no-walrasiano.

Los desempleados no serán capaces de cumplir con los requerimientos de colaterales bancarios. Su riqueza está por debajo de K_b^*, el mínimo valor de un colateral. Además, los costos de transacción (selección, evaluación, monitoreo) son mayores cuando los bancos otorgan créditos a pequeñas firmas, y los trabajadores solo pueden implementar firmas pequeñas. Debido a que existen economías de escala en la oferta de crédito bancario, no puede existir un mercado de crédito al por menor, y ellos quedan fuera de juego. Por lo tanto, el banco considerará no rentable otorgar pequeños préstamos porque su costo unitario es alto. A mayor tamaño del préstamo, mayores los beneficios del banco. Los bancos encuentran más rentable otorgar un pequeño número de préstamos grandes, que un gran número de préstamos pequeños. Luego, existe exclusión de prestamistas potenciales de pequeños montos de préstamos.

Otra condición para establecer una firma capitalista y ser capaz de mantenerla es proteger el capital físico contra el ries-

go de destrucción. Como en el caso del crédito bancario, las pequeñas unidades de producción no pueden ser clientes de una compañía de seguros. Los costos de transacción en ese caso son también mayores para las unidades pequeñas. Como en el caso del crédito bancario, existen economías de escala en la oferta de seguro que hacen imposible el mercado al menudeo. Por lo tanto, el costo unitario con firmas pequeñas es más elevado y las compañías de seguro no encontrarán rentable vender servicios de seguros a estas firmas pequeñas. Luego, los desempleados no podrán diseminar su riesgo a través del mercado de seguros; es decir, no podrán reducir sus riesgos por medio de la compra de seguros en el mercado de seguros. Los trabajadores que hubiesen sido capaces de establecer una firma estarán, por lo tanto, sujetos a riesgos individuales no asegurados.

En suma, en la economía épsilon, los capitalistas usan el desempleo como un instrumento de disciplina para los trabajadores. Por ello, algunos trabajadores son excluidos del mercado de trabajo. Dado que estos trabajadores están también excluidos del mercado de créditos y de seguros, no podrán escapar del desempleo.

¿Cuál es el papel del Gobierno? Si se otorgara un seguro de desempleo a los desempleados, equivalente al salario de mercado, este pago haría que dejasen de buscar empleo y la disciplina laboral se perdería. El mercado de trabajo se convertiría en un mercado walrasiano. Por lo tanto, el Gobierno tendría que pagar solo una fracción de la tasa de salario de mercado para conservar el sistema de incentivos de disciplina laboral. El seguro de desempleo mantendría así el exceso de oferta laboral. Es esta política, en efecto, la que se observa en el primer mundo.

El Gobierno podría también actuar como empleador de última instancia. Pero esta política de pleno empleo haría del sistema capitalista inoperante y sería resistida por los capitalitas. Los capitalistas podrían bloquear el intento del Gobierno ame-

nazando reducir la inversión si se siguiese una política de pleno empleo. La situación de pleno empleo, la eliminación del desempleo como instrumento de disciplina laboral, constituye un desafío a la posición social de privilegio de los capitalistas, aun cuando los beneficios pudiesen aumentar con dicha política. Pero sus preferencias están jerarquizadas y la posición social es prioritaria. Este conflicto de intereses entre el Gobierno y la clase capitalista, donde la última puede finalmente imponer el poder del capital, puede ser denominado como la «amenaza kaleckiana».[3] Detrás de este comportamiento yace la teoría presentada más arriba: el desempleo es necesario para el funcionamiento de la sociedad épsilon.

El poder del capital significa que los capitalistas pueden imponer límites al nivel de desempleo en la economía. Pueden poner también límites a la forma en la que el seguro de desempleo es construido. La compensación a los desempleados debe ser menor que la tasa salarial de mercado, para mantener el principio de la disciplina laboral. La amenaza kaleckiana puede ser extendida a cualquier acción de gobierno que les cree conflicto de intereses.

La economía épsilon también opera con desigualdad persistente. El ingreso nacional se distribuye entre beneficios y salarios; el ingreso medio de los capitalistas es mucho mayor que el ingreso medio de los trabajadores. Los más pobres serán los desempleados. Esta inequidad persistirá debido a la exclusión de los trabajadores de los mercados de crédito y seguros. El funcionamiento del mercado de créditos restringe el acceso de

[3] Kalecki dijo lo siguiente: «[la inversión privada] le da a los capitalistas un poderoso control indirecto sobre la política de gobierno: todo aquello que pudiese sacudir el estado de confianza debe ser cuidadosamente evitado porque puede causar una crisis económica» (Kalecki 1943: 139). El mismo autor señaló: «La disciplina en las fábricas y la estabilidad política son más apreciadas que los beneficios por los líderes de los negocios. Su instinto de clase les dice que el pleno empleo duradero es un error desde su punto de vista y que el desempleo es una parte integral del sistema capitalista normal» (1943: 141).

los trabajadores al capital, mientras que el mercado de seguros se rehúsa a protegerlos ante los riesgos.

Lucas (1992) desarrolló un modelo teórico para mostrar el rol de los riesgos en la generación de la desigualdad. Su argumento básico es el siguiente. Considérese una sociedad en la cual las familias tienen como dotaciones un flujo permanente de ingresos que es igual para todos; donde las familias tienen iguales preferencias intertemporales de consumo; y donde las familias estén sujetas a riesgos individuales no asegurados. En esta sociedad, la desigualdad aparecerá en el largo plazo y crecerá indefinidamente. Empezando con una posición de completa igualdad, la sociedad experimentará incrementos continuos de desigualdad debido a *shocks* negativos contra las dotaciones.

Si inicialmente los individuos están dotados de cantidades de capital desiguales, el rol del riesgo en la generación de desigualdad es mucho más evidente. Los ricos serán capaces de jugar juegos con altos retornos y altos riesgos porque su capacidad de absorber riesgos es alta. Si resultan ganancias, su riqueza aumentará y su capacidad de absorber riesgos aumentará también, y jugarán juegos aún más riesgosos que les ofrecerán la posibilidad de mayores ganancias aun, y así sucesivamente. Los pobres no tienen capacidad económica para hacer esto y permanecerán siendo pobres.

Desde el punto de vista de la reproducción de la desigualdad, los mercados del trabajo, el crédito y del seguro juegan un rol significativo. En este estudio se les llama, por eso, *mercados básicos*. En la economía épsilon, existe una jerarquía de mercados. La diferencia con la teoría neoclásica es notoria, pues allí todos los mercados son de igual importancia. En resumen, la persistencia de la desigualdad se debe al hecho de que los trabajadores son excluidos de los mercados básicos.

La desigualdad en el ingreso no puede hacerse más desigual porque cambios en las variables exógenas tienen efectos ambiguos. No puede hacerse menos desigual porque existen ciertos

mecanismos de exclusión en el funcionamiento del sistema que impiden dicha reducción.

Como se demostró anteriormente, existen mecanismos de exclusión en el proceso económico que conducen a la persistencia de esta desigualdad. Los más significativos son los siguientes:

a. La dotación inicial de capital está distribuida de forma desigual. Los trabajadores son excluidos de la propiedad. Los beneficios son apropiados por la clase capitalista y son usados para aumentar aún más la propiedad.
b. Los más ricos pueden participar en aquellos juegos económicos que poseen altos rendimientos pero también altos riesgos, debido a que tienen capacidad económica para absorber esos riesgos si llegasen a ocurrir. Debido a sus bajos ingresos, los trabajadores son excluidos de dichos juegos.
c. Dado que el desempleo es necesario para la disciplina laboral y los beneficios, habrá trabajadores que serán excluidos del mercado de trabajo y se convertirán en desempleados.
d. Los más ricos tienen acceso a crédito bancario y al mercado de seguros. Los trabajadores están excluidos de estos mercados.
e. Los capitalistas pueden generar innovaciones ahorradoras de trabajo. Los trabajadores son excluidos del proceso de generación de innovaciones tecnológicas.

Todas estas proposiciones se derivan del modelo dinámico de la economía épsilon. Por lo tanto, la persistencia de la desigualdad —su viscosidad— es endógena.

En la sociedad omega, la desigualdad inicial también tiende a reproducirse. Debido a que el sector capitalista de omega opera como en épsilon en cuanto a la acumulación de capital y a que el sector de subsistencia es residual, el equilibrio dinámico de la desigualdad tendrá las mismas propiedades que la de épsilon. Para la sociedad sigma se aplican, a su vez, las conclu-

siones obtenidas en omega. La desigualdad inicial también tiende a reproducirse y puede incluso aumentar debido a que la exclusión social impone barreras adicionales a la movilidad social de los más pobres. En suma, las conclusiones que se han obtenido para la sociedad épsilon sobre la reproducción de la desigualdad inicial en el equilibrio dinámico (cuando la inversión privada es endógena) se pueden extender a las sociedades omega y sigma. En estas tres sociedades, existen barreras económicas que impiden la reducción de la desigualdad; en sigma, en adición, existen barreras sociales.

En cada sociedad, la desigualdad en la distribución de los flujos de ingresos sigue dependiendo únicamente de la desigualdad inicial en la dotación de los *stocks* de activos económicos y sociales. Modificar la desigualdad inicial, es decir, su legado, implica aplicar *shocks* refundacionales a las sociedades. Estos *shocks* no se han dado dentro de países capitalistas. Así se puede entender la constancia de la desigualdad por países, uno de los hechos estilizados de la economía mundial que los estudios empíricos han demostrado (cf. Li, Squire y Zou 1998).

Queda pendiente ahora la tarea de explicar el otro hecho estilizado de la economía mundial: la disparidad persistente en el nivel del ingreso per cápita entre los países del primer mundo y tercer mundo. Un primer paso consiste en comprender la lógica de los inversionistas en cuanto a la asignación de su capital a los distintos países de la economía mundial. A esa tarea se dedica el capítulo siguiente.

Capítulo 8
Bienes públicos e inversión privada

No existe una teoría económica separada sobre la distribución del ingreso en las teorías de equilibrio general que han sido estudiadas hasta ahora —épsilon, omega y sigma—. La producción y la distribución son determinadas simultáneamente en un único proceso económico. La desigualdad del ingreso es entonces un resultado del proceso económico.

Una de las características de la distribución del ingreso generada por estas teorías de equilibrio general es que el grado de desigualdad del ingreso puede tomar cualquier valor. Cualquier grado de desigualdad es socialmente viable, es decir, que las sociedades pueden tolerar cualquier grado de desigualdad. No hay límites a la desigualdad.

Ciertamente, cada una de las sociedades estudiadas opera bajo restricciones. En la sociedad épsilon existe un piso para el salario real y la tasa de desempleo; en las sociedades omega y sigma, el salario real relativo a los ingresos del sector de subsistencia tiene límites. Las sociedades épsilon, omega y sigma imponen restricciones a la caída del salario nominal. Pero estas son restricciones sobre las tasas salariales, no sobre el resultado de la distribución del ingreso. En todas estas teorías, el grado de desigualdad en el ingreso puede tomar cualquier valor en el equilibrio general. Por lo tanto, una vez que el equilibrio general es alcanzado, la producción y distribución se repiten periodo tras periodo bajo las mismas reglas del juego económico; es decir, el equilibrio general implica orden social.

Esta predicción es inconsistente con los hechos más distintivos acerca del funcionamiento de los países capitalistas. Los trabajadores realizan huelgas, los pobres expropian los activos

de otras personas usando medidas violentas, tales como robos, secuestros, invasiones a la propiedad privada. Los Gobiernos gastan proporciones importantes de sus presupuestos en medidas de protección a la propiedad y dirigidas al orden social —tales como las fuerzas policiales, la administración de la justicia y las cárceles—. Estos gastos, que pretenden restaurar el orden social, son financiados por los impuestos. Las personas también realizan gastos directos para proteger sus activos, y así se desarrolla la industria de seguridad. Todos estos recursos utilizados en la protección de los derechos de propiedad no son insignificantes. ¿Cómo se puede explicar la conducta humana que crea el desorden social? Creando una teoría, por supuesto.

8.1. Tolerancia social a la desigualdad

La idea básica es que el orden social no puede existir independientemente del grado de desigualdad de la sociedad. Con el fin de desarrollar esta idea en la forma de teoría, existe la necesidad de una teoría acerca de la racionalidad en las personas sobre la desigualdad. Tal teoría será presentada aquí. Se busca que esta teoría sea válida en todas las sociedades capitalistas que hemos estudiado. Esta teoría puede ser llamada *teoría de la tolerancia limitada a la desigualdad.*

La siguiente proposición alfa deberá ser añadida a aquellas que fueron establecidas para las sociedades épsilon, omega y sigma:

$\alpha_{\varepsilon,\omega,\sigma}(5)$. *Tolerancia social limitada a la desigualdad.* Los individuos tienen un sentido de justicia o equidad respecto a la desigualdad económica; es decir, los individuos tienen una tolerancia limitada a la desigualdad. Si la desigualdad alcanza un grado superior al que indica el umbral de tolerancia, la gente reaccionará desafiando las reglas del contexto institucional.

Un modelo particular de esta teoría se presenta en este capítulo. Considere la siguiente función de utilidad para el individuo *j:*

$$U_j = U_j(Y_j, Y_j / Y_k), \tag{8.1}$$

$$\text{tal que } U_j > 0 \text{ si } (Y_j, R_j) \geq (Y_j^*, R_j^*)$$
$$U_j < 0 \text{ si } (Y_j, R_j) < (Y_j^*, R_j^*)$$

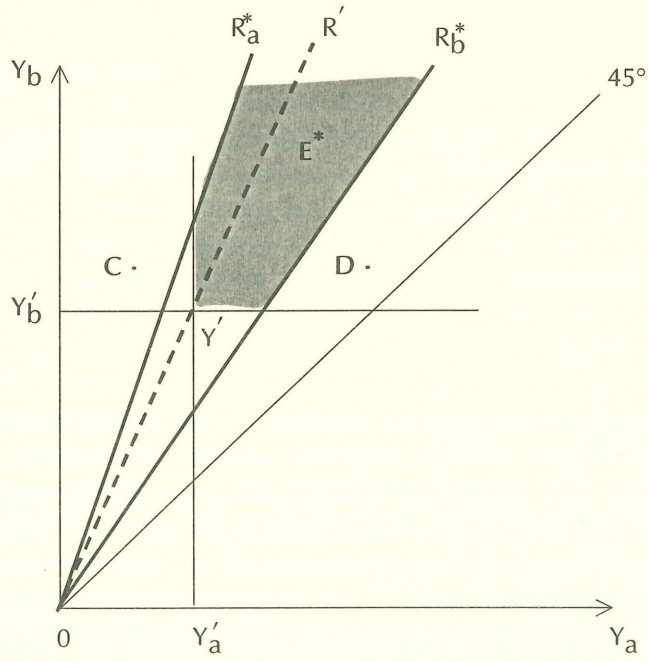
Aquí, Y_j y $R_j = Y_j/Y_k$ indican el ingreso del individuo *j* en forma absoluta y relativa, siendo *k* su grupo de referencia. Los valores marcados con asterisco representan los valores del umbral de tolerancia a la desigualdad.

Esta función de utilidad supone que el individuo considera a ambos tipos de ingresos como si fuesen bienes y, por lo tanto, prefiere más de ambos a menos de ambos. También supone que el individuo está dispuesto a renunciar a una parte de su ingreso absoluto a cambio de un ingreso relativo mayor. Adicionalmente a las mercaderías, el individuo también busca consumir estatus social. La función de utilidad supone que los individuos tienen un sentido de justicia en lo que respecta a los ingresos relativos y absolutos. El valor negativo que el individuo le atribuye a su utilidad cuando cualquiera de los dos ingresos sobrepasa el umbral de tolerancia significa que esta situación es intolerable y que esta persona tomará acciones como respuesta. Esta situación causa en el individuo no solamente una pena subjetiva (un caso de externalidad en su función de utilidad), sino que la persona toma acciones para resistir este resultado.

Esta lógica económica es consistente con la teoría de Maslow, pues se incorpora las necesidades primarias de la gente, aquellas asociadas a la supervivencia física y social; es decir, se supone que la gente no acepta pasivamente situaciones extremas de pobreza relativa y absoluta. Los valores de los umbrales en la función de utilidad son exógenamente determinados y son inobservables.

Esta teoría se ilustra en la Figura 8.1. Considérese una sociedad compuesta por dos grupos sociales, A y B. Supóngase que la función de utilidad es igual para todos los individuos dentro de cada grupo. Supóngase, además, que la distribución del ingreso obtenido del equilibrio general es Y'= (Ya', Yb'), donde el grupo A es el relativamente pobre. Para simplificar aún más, supóngase que los umbrales de ingreso absoluto son los ingresos corrientes generados por el equilibrio general; es decir, los individuos no toleran una caída en su ingreso real. Es evidente que cada rayo que sale del origen mide un ingreso relativo y, por lo tanto, mide también un mismo grado de desigualdad. Supóngase, también, que los umbrales de tolerancia para el ingreso relativo del grupo A es R_a^* y para el grupo B es igual a R_b^*.

Figura 8.1
Región de la tolerancia social a la desigualdad



Con estos datos, el conjunto de ingresos que son tolerados por ambos grupos es la región sombreada E^*. Este conjunto será llamado *región de tolerancia social.* Este es un subconjunto estricto de todas las posibles distribuciones. El conjunto E^* es socialmente tolerable porque esa desigualdad es aceptable por todos los individuos.[1]

La predicción de esta teoría es que el desorden social será mayor cuanto mayor sea el grado de desigualdad. Para unos umbrales de tolerancia dados y para un grado de desigualdad dado, habrá un grupo de individuos que considera esa desigualdad inaceptable, y reaccionará y creará desorden social. Si el grado de desigualdad aumenta, el grupo de individuos que considera que esa desigualdad es inaceptable será de un mayor tamaño, pues al grupo inicial se unirán otros que consideraban el anterior grado de desigualdad aceptable, pero no el de ahora. Y así sucesivamente. El algoritmo genera la proposición beta indicada.

En una situación de equilibrio general, existirá una tolerancia social a la desigualdad en la sociedad si y solo si la distribución cae en la región de tolerancia social. Debido a que, en este caso, la distribución del ingreso es un resultado de la solución del equilibrio general y se encuentra en la región de la tolerancia social, existirá equilibrio general con orden social en esta economía. Si la solución de equilibrio general es tal que la distribución del ingreso cae fuera de la región de tolerancia social, existirá equilibrio general pero sin orden social. Este grado de desigualdad no será socialmente aceptado, y aquellos que no estén contentos con esta situación propiciarán revueltas, es decir, ellos cambiarán su conducta. Los perdedores tratarán de recuperar la situación de desigualdad tolerable a través de una

[1] Dentro de una firma, una excesiva desigualdad entre los trabajadores (una distribución de ingresos fuera de E^*) generará ineficiencia económica porque esta desigualdad perjudica el trabajo en equipo y la cooperación. Entonces, las firmas buscarán un grado de desigualdad que sea tolerable, como parte de su lógica de maximización de beneficio.

redistribución privada y violenta; ellos se rebelarán en contra de los derechos de propiedad privada. Por ejemplo, un trabajador que pierde su empleo, y no es capaz de encontrar otro, podría convertirse en un delincuente. El grado de violencia en la sociedad es, entonces, endógeno.

No hay nada en las teorías de equilibrio general que prevenga que la desigualdad del ingreso caiga fuera de la región de tolerancia social. Tampoco existen mecanismos automáticos que puedan restaurar la desigualdad de una situación de intolerancia a otra de tolerancia social. La función de utilidad que indica la tolerancia a la desigualdad que tienen los individuos no está internalizada en algún lugar dentro del intercambio de mercado. La región de tolerancia social a la desigualdad no es parte de las restricciones en cualquier teoría de equilibrio general. En este sentido, uno podría decir que la desigualdad en el ingreso no está autorregulada. Entonces, es perfectamente posible que el equilibrio general pueda generar un grado de desigualdad que esté fuera de la región de tolerancia social, como en C en la Figura 8.1. Este resultado genera desorden social.

Cada vez que se toca el tema de la equidad en la distribución del ingreso se invoca el trabajo clásico del filósofo John Rawls (1971). ¿Cómo se compara esta teoría con la de Rawls? Su teoría puede ser transformada en teoría económica en los siguientes términos. En una sociedad donde los individuos enfrentan el mismo conjunto de circunstancias, donde todos pueden ponerse en la situación de los otros, la regla de consenso sobre la distribución será: (a) Todos estarán de acuerdo con un ingreso mínimo garantizado, porque todos están sujetos al riesgo de no percibir ingresos, es decir, todos estarán de acuerdo con que todos estén protegidos contra esa posibilidad; (b) Ellos estarán de acuerdo en imponer límites a la desigualdad, pues, debido a que todos son proclives a obtener ganancias extraordinarias, todos estarán de acuerdo en compartir colectivamente parte de ese ingreso.

En este contexto, cualquier individuo guiado por su propio interés diseñaría esta regla distributiva independientemente. Esta es una regla impersonal. Es simétrica. Es justa. Es parte del contrato social. En la Figura 8.1, las reglas distributivas de Rawls generan una región de tolerancia social que es simétrica con respecto a la línea de 45 grados. Esta es una teoría diferente a la que se ha presentado aquí.

Claramente, el mundo abstracto de Rawls es uno en el cual no existen estructuras o clases sociales. Esta es una regla para todas las sociedades flexibles y abiertas, en las que la movilidad social no está limitada por barreras institucionales (donde *la tortilla puede darse la vuelta*). En esta sociedad, la probabilidad de lograr una ubicación en el estrato alto de la distribución del ingreso es la misma para cada uno; de manera simétrica, la probabilidad para ubicarse en el estrato bajo también es la misma para todos. Las personas juegan con dados que no están cargados. En una sociedad de clases, por contraste, algunas personas juegan con los dados cargados. Las oportunidades de ubicarse en ciertos estratos de la pirámide de ingresos son diferentes entre las personas. Este hecho se debe a la existencia de mecanismos de exclusión que caracterizan a todas las formas de capitalismo.[2]

El economista Albert Hirschman (1973) escribió un artículo sobre la tolerancia a la desigualdad que también es un clásico en la literatura. Su proposición es que la gente tolera la des-

[2] La misma idea de contrato social parece requerir una sociedad liberal, en la que la libertad individual esté garantizada. El contrato social no puede ser aplicable como un concepto económico en una sociedad feudal o esclavista. Por extensión, no puede ser aplicable a una sociedad de clases. Un ejemplo simple que ilustra el tipo de contrato social de Rawls está relacionado con las reglas de uso de las calles para autos o peatones. En el primer mundo, la regla es que los peatones tienen la preferencia sobre los autos. Aquí todos tienen la misma oportunidad de poseer un carro. Todos se pueden poner entonces en los zapatos de los otros. La regla es justa. En el tercer mundo, solo una pequeña fracción de la población tiene la oportunidad de poseer un carro, las demás personas serán solo peatones. La regla es que el carro tiene preferencia sobre los peatones. Esta regla es una imposición social de los ricos.

igualdad solo cuando es transitoria. Por ejemplo, en las etapas iniciales del crecimiento económico, cuando la desigualdad se está incrementando, la gente tolerará esta desigualdad si sus expectativas son que esas inequidades se reducirán más tarde. Si esas expectativas se incumplen, se producirá un desorden social. Hirschman utilizó como metáfora de su teoría el embotellamiento de vehículos en un túnel, en el que los vehículos empiezan a moverse, excepto en la fila de un grupo.

La teoría de Hirschman tiene el problema de que introduce las expectativas, una variable inobservable. Para que su teoría sea refutable, las expectativas de la gente se tienen que suponer dadas. Luego, cuanto más pasa el tiempo, el desorden social será mayor pues habrá cada vez más gente que no ve cumplidas sus expectativas. Es pues una teoría de corto plazo sobre la trayectoria en el tiempo que describe el desorden social para un aumento dado en la desigualdad. En la teoría presentada aquí se trata de la tolerancia de la gente a grados de desigualdad cada vez mayores. Así, es claro que la teoría de Hirschman es complementaria, pues añade la dimensión temporal a la generación del desorden social que resulta de un cambio en el grado de desigualdad.

8.2. Equilibrio general con desorden social

Cuando la distribución del ingreso cae en la región de tolerancia social, la economía consigue el equilibrio general con orden social. Entonces, en un sistema estático, el proceso económico será repetido periodo tras periodo, con la misma producción y la misma desigualdad del ingreso, hasta que ocurra un cambio en una variable exógena. Implícitamente, estos son los tipos de resultados que han sido generados por las soluciones de equilibrio general estudiadas en las economías épsilon, omega y sigma.

Si la distribución del ingreso cae fuera de la región de tolerancia social, existirá un equilibrio general con desorden social. Los

precios y cantidades estarán determinados por el sistema de mercado, pero los perdedores actuarán por fuera del mercado y darán lugar a *shocks* negativos sobre los ingresos o los activos de capital de los ganadores. Los perdedores tomarán acciones tratando de redistribuir flujos y *stocks* por la fuerza o por medios no contractuales. La regla de los derechos de la propiedad privada no será plenamente respetada. Los agentes internalizarán la situación y cambiarán su forma de producción e intercambio. El equilibrio general resultante tendrá los valores de equilibrio en la producción y distribución, junto con la redistribución privada y violenta (un tipo de impuesto privado). Mientras que las variables exógenas del sistema permanezcan constantes, estos valores de equilibrio se repetirán periodo tras periodo.

El orden social es un bien público. Es indivisible; y, una vez que es producido, nadie puede ser excluido de su beneficio. De esta forma, los ganadores no tienen incentivos de solucionar este problema de manera individual. Y la acción colectiva fracasaría por el problema olsoniano; es decir, no habrá un mecanismo espontáneo que mueva el equilibrio del punto C a un punto dentro de la región E^*.

En este caso, y contrariamente a la proposición de Adam Smith, los individuos actúan guiados por su propio interés en el mercado pero son conducidos como por una mano invisible no al bien común sino a una situación de excesiva desigualdad y desorden social. Demasiada desigualdad puede ser un resultado no buscado por los individuos, un claro caso de falla de mercado. La desigualdad moderada no genera ineficiencias en la economía, pero la desigualdad excesiva sí las genera.

El Gobierno podría llevar a cabo la redistribución necesaria. Sería una redistribución hecha de forma ordenada y podría incluir los *stocks* de activos o los flujos de ingresos, o una combinación. Gracias a la intervención del Gobierno, la sociedad puede retornar al equilibrio general con orden social. En la Figura 8.1, la distribución del ingreso C —un resultado del equilibrio general— puede ser trasladado a la región E^*. Si la

redistribución del Gobierno está dirigida a una distribución del ingreso como la situación D, esta tampoco será socialmente tolerable, pues ahora existe demasiada igualdad. Pero, ¿tiene el Gobierno los incentivos para llevar a cabo esta redistribución?

Si la acción del Gobierno no existe o es insuficiente para restaurar la distribución del ingreso a la región de tolerancia, entonces la pugna por la redistribución conducirá a la violencia. Los individuos amenazados con la expropiación de la propiedad por medios privados buscarán asegurarse de estos riesgos. Este incremento en la demanda por seguros propiciará la expansión de la industria de seguros. Gracias a esa solución de mercado, la sociedad podrá retornar a un equilibrio general con orden social.

La solución de mercado no será suficiente, pues hay riesgos que no se pueden asegurar. Las personas que son amenazadas con la redistribución privada buscarán medidas para actividades de seguridad propia. Esta demanda propiciará la oferta y, entonces, existirá una expansión de la industria de la seguridad (producción de bienes y servicios para la protección de la propiedad privada y la vida).

El equilibrio general con un grado de desigualdad fuera de la región de tolerancia social será una situación en la cual la sociedad asigna sus escasos recursos a la protección de la propiedad privada. Esta asignación puede ser hecha a través del Gobierno y por los propios individuos; esto es, a través de la oferta de bienes públicos y bienes privados. El uso de estos recursos es una medida del costo económico de la excesiva desigualdad. En esta situación existirá violencia en la sociedad. Los perdedores buscarán la redistribución del ingreso hasta reponer la distribución del ingreso a la región de tolerancia. Esto se hará a través de acciones colectivas e individuales, legales e ilegales. Si el Estado falla en responder a estas demandas, sea por acción u omisión, el desorden social será mayor. El desorden social es, entonces, el resultado de una falla del mercado y del Gobierno a la vez.

8.3. Proposiciones beta

Los umbrales de tolerancia no son observables. Para que la teoría sea falsable, se tiene que suponer que estos umbrales se mantienen fijos. Si el ingreso absoluto o el ingreso relativo de un grupo social dado cae continuamente, entonces más gente habrá sobrepasado los valores de sus umbrales de tolerancia y más gente reaccionará. Por lo tanto, la teoría de la tolerancia limitada a la desigualdad predice que el grado de orden social depende del grado de desigualdad del ingreso. Debido a que el orden social es un insumo en el proceso económico, sociedades altamente desiguales funcionarán de manera diferente a sociedades más igualitarias.

Del modelo teórico propuesto se pueden derivar las siguientes proposiciones beta sobre el comportamiento global de las sociedades:

(a) Cuanto mayor sea el grado de desigualdad, mayor el grado de desorden social para todo tipo de sociedad capitalista.
(b) Para dos sociedades capitalistas, con iguales umbrales de tolerancia a la desigualdad, la economía con mayor grado de desigualdad mostrará un mayor grado de desorden social.
La teoría de la tolerancia a la desigualdad también predice que, en una economía capitalista de cualquier tipo, cuanto *mayor la disminución del ingreso absoluto o relativo de los trabajadores:*
(c) Mayor será el nivel de violencia en la sociedad y, además, mayor será el monto de recursos que la sociedad usará para proteger la propiedad privada; y, entonces, la industria de la seguridad será más grande. Estos efectos dependerán de la redistribución hecha a través del Gobierno. La predicción se refiere a la violencia individual. La teoría no predice la violencia colectiva porque esta requeriría de la acción colectiva de los agentes, cuya materialización depende de factores exógenos.

(d) Más gente se rebelará ante el sistema político y económico. Las actividades de oposición política se incrementarán. El conflicto político se intensificará. La inestabilidad política crecerá. La inestabilidad política es entonces endógena.
(e) En el corto plazo, la productividad del trabajo caerá. Más gente tratará de suplir la pérdida en el ingreso absoluto o relativo vía ingresos no contractuales. Los costos marginales de las firmas se incrementarán y el producto total caerá. Otra forma de expresar esta idea es introduciendo el orden social como un factor de producción, como un bien público, en la función de producción de la firma.
(f) En el largo plazo, la inversión privada caerá. Los inversionistas esperarán que las políticas del Gobierno intenten restaurar un grado socialmente aceptable de desigualdad. Esta política conducirá a mayores tasas impositivas sobre los ricos. Esta medida reducirá los beneficios y la tasa de retorno de las inversiones. Los capitalistas podrían, además, sentir que los trabajadores quieren expropiar sus propiedades. Los capitalistas podrían esperar que la inestabilidad política aumentase, y eso incrementaría los riesgos de cambios en las reglas del juego. En cualquier caso, las amenazas de que se incrementen los impuestos, la expropiación, o que las reglas sean cambiadas son todos factores de riesgo que los inversionistas toman en cuenta antes que ellos realicen sus inversiones en esta sociedad.

Mientras que las proposiciones beta (a)-(d) son evidentes por sí mismas, (e) y (f) necesitan verificaciones lógicas. Las siguientes secciones desarrollan estas cuestiones.

8.4. Bienes públicos en la función de producción

Considérese una función de producción agregada de la siguiente forma:

$$Q_b = f(D_h, K_b, Z) \qquad (8.2)$$
$$= f(L, K_b, K^i_b, K_h, S)$$

La primera ecuación indica que el producto total Q_b depende de la cantidad de trabajadores D_h, del *stock* de capital físico de las firmas K_b y del *stock* de bienes públicos Z. La segunda ecuación desagrega el vector de bienes públicos Z en tres componentes: infraestructura física (K^i_b), el *stock* de capital humano (K_h) y el grado de orden social (S). La cantidad de trabajo es ahora medida como número de personas (L), pues el capital humano aparece separado. Un incremento del capital humano promedio, manteniendo la misma cantidad de personas, incrementa el nivel de producto. El capital humano es incrementado a través del gasto público, como es el caso de la infraestructura física.

Considérese, ahora, el otro tipo de bienes públicos que está incluido en el vector Z: el orden social. Claramente, el orden social es un bien público: una vez que está presente, nadie puede ser excluido de sus beneficios. Como un bien público, el grado de orden social se convierte en un factor de la función de producción. Dado el mismo conjunto de insumos, el producto será más pequeño cuando el grado de orden social es menor. Algunas cantidades de insumos destinados a proteger la propiedad privada (defensas y sistemas de seguridad) y a pagar sobornos son también añadidos al proceso de producción, aun cuando ellos no sean necesarios desde un punto de vista tecnológico. Entonces, aparece la ineficiencia en el proceso productivo.

¿Qué factores determinan el orden social? A partir de la teoría de la tolerancia a la desigualdad, el orden social depende del grado de desigualdad económica. Un mayor grado de desigualdad implica un grado mayor de desorden social.

El grado de desigualdad en la distribución de los flujos de ingreso de una sociedad resulta del proceso económico. Esta es una variable endógena. Pero esta desigualdad depende de la

desigualdad inicial en los *stocks*, es decir, la desigualdad inicial en las dotaciones de activos económicos y sociales entre los individuos. Como se ha mostrado en los capítulos anteriores, los otros factores exógenos no cambian mucho la desigualdad inicial. Por lo tanto, se puede concluir que la desigualdad del ingreso depende de la desigualdad inicial en los *stock* de activos (δ_0). Este resultado implica que las sociedades que «nacieron» más desiguales al capitalismo, continuarán siendo así, y sociedades que «nacieron» menos desiguales al capitalismo también continuarán siendo así. La desigualdad es un rasgo permanente de las sociedades, una característica estructural. Por lo tanto, el grado de desorden social es también una característica estructural de las sociedades.

Por lo tanto, la función de producción agregada de la economía puede ser escrita como:

$$Q_b = F (L, K_b, K_b^i, K_h, \delta_0), F_5 <0 \text{ y } F_i >0 \qquad (8.3)$$

En consecuencia, en el corto plazo, la productividad del trabajo dependería también del grado de desigualdad estructural de la sociedad. Manteniendo las otras cosas iguales, las sociedades más desiguales serán menos productivas.

8.5. Bienes públicos en la inversión privada

En el largo plazo, los insumos varían en cantidad y calidad. Capital físico y capital humano son acumulados, y la tecnología mejora. En el sistema dinámico, la inversión privada en capital físico es endógena. En los modelos dinámicos del capítulo anterior, la inversión dependía positivamente de las ganancias. Los bienes públicos no jugaban papel alguno, ni el grado de desigualdad que subyace al grado de orden social. La desigualdad es ahora introducida en la función de inversión.

Debido a que la decisión de inversión se realiza bajo condiciones de incertidumbre, ella depende del retorno esperado y de la capacidad de los inversionistas de absorber los riesgos. Cuanto mayor sea esta capacidad, los inversionistas entrarán a proyectos de inversión con retornos esperados mayores, aun cuando estos sean más riesgosos. Esta capacidad está determinada por la cantidad de riqueza que el inversionista tiene. El capitalista determina la elección de su portafolio de manera jerarquizada: primero por sus restricciones en cuanto a riqueza y, segundo, por sus preferencias. Aquí se supone que los inversionistas escogen sus inversiones de acuerdo con la media y varianza de aquellos proyectos que son parte de su conjunto factible —conjunto que excluye proyectos que impliquen pérdidas posibles que lo llevarían al desastre económico— (ver el apéndice para una presentación formal de esta teoría de la conducta frente al riesgo).

Se supone, además, que el capital es completamente móvil a través de las economías. Pero, ¿cómo deciden los inversionistas la asignación de sus fondos de inversiones a esas distintas economías? Se presenta ahora un modelo teórico que intenta responder esta pregunta. El supuesto es que el retorno esperado de una inversión privada en una sociedad dada, así como la varianza de ese retorno, dependen del *stock* de bienes públicos en el país (Z). Mientras mayor sea el *stock* de bienes públicos, mayor será el retorno medio esperado de la inversión y menor será el riesgo; por lo tanto, mayor será la inversión. El otro supuesto es que el *stock* de bienes públicos tiene tres componentes: infraestructura física (K_b^i), capital humano (K_h) y orden social (S). Se ha mostrado que el orden social depende negativamente del grado de desigualdad en la distribución de ingresos; dicho grado, a su vez, depende de la desigualdad inicial en la dotación de activos económicos y sociales (δ_0).

Los supuestos sobre la inversión privada pueden ser representados por el siguiente conjunto de ecuaciones:

$$
\begin{aligned}
&K_b(t+1) = K_b(t) + I_b(t) \\
&I_b(t) = \alpha\, P(t) + FI(t),\ 0 \le \alpha \le 1 \\
&\alpha = f(Z(t)) = f(K_b^i,\ K_h,\ \delta_o),\ f_1 < 0,\ f_2 > 0,\ f_3 > 0 \\
&FI(t) = g(Z(t)) = g(K_b^i, K_h, \delta_0),\ \ g_1 > 0, g_2 > 0, g_3 < 0 \\
&I_b(t) = \psi(P(t), K_b^i(t), K_h(t), \delta_0)
\end{aligned}
\qquad (8.4)
$$

La primera ecuación representa la identidad de la acumulación del capital. La segunda ecuación nos dice que la inversión privada no depende solo de los ahorros domésticos. Tanto la proporción α del total de los beneficios que es reinvertida localmente —hecho que implica que la proporción $(1-\alpha)$ es invertida fuera— como el influjo de inversión extranjera (FI) dependen del *stock* de bienes públicos de la economía. La última ecuación muestra la forma reducida.

Pero el capital de infraestructura y el capital humano pueden también ser acumulados a través de la inversión. Entonces, se puede escribir:

$$
\begin{aligned}
&K_b^i(t+1) = K_b^i(t) + I^i(t) \\
&I^i(t) = \theta(Y(t), \delta_0),\ \theta_1 > 0,\ \theta_2 < 0 \\
&K_h(t+1) = K_h(t) + I_h(t) \\
&I_h(t) = \phi(Y(t), \delta_0), \phi_1 > 0, \phi_2 < 0
\end{aligned}
\qquad (8.5)
$$

La inversión en esos dos tipos de capital depende de la conducta del Gobierno, que depende de sus ingresos tributarios —el que a su vez depende del ingreso nacional— y de la desigualdad inicial de la sociedad. Sociedades más desiguales (con mayor grado de exclusión social) darán lugar a asignaciones del presupuesto público menos favorables al capital humano.

En la forma reducida, la función de la inversión privada deviene en:

$$I_b^{\circ}(t) = \psi(K_{bO}, K_{bo}^{i}, K_{ho}, \delta_0; S_h, S_m, z^*, r^*; t) \qquad (8.6)$$

Esta ecuación muestra el equilibrio dinámico de la inversión privada. La trayectoria en el tiempo es creciente, manteniendo fijas las variables exógenas. La dinámica comparativa muestra que todas las variables exógenas tienen efecto positivo sobre la trayectoria (la desplazan hacia arriba), excepto la desigualdad inicial δ_0 y la tasa de interés internacional r^*, cuyos efectos son negativos.

La desigualdad inicial juega un papel significativo en la inversión y, por lo tanto, en el crecimiento económico de la sociedad. El efecto de la desigualdad vía beneficios es negativo, porque la mayor desigualdad (y el mayor desorden social que implica) afecta negativamente la productividad de corto plazo. El efecto vía capital humano es negativo porque mayores desigualdades iniciales implican menos derechos para la gente y menor presión del grupo para obtener mayor calidad de educación pública. El efecto vía el orden social es negativo porque una mayor desigualdad inicial implica una sociedad con más desorden social. La sociedad es más inestable y más riesgosa para la inversión.

Este resultado puede, además, ser presentado en términos de las regiones de la tolerancia social. Si el grado de desigualdad se encuentra en la región de tolerancia social, entonces la inversión no es afectada por cambios pequeños en la distribución del ingreso (este era el supuesto implícito hecho en el modelo dinámico de la economía épsilon). Sin embargo, si el grado de igualdad se encuentra fuera de esta región, entonces el nivel de inversión caerá. Así, la inversión privada no es independiente del grado de desigualdad de la sociedad.

Una condición suficiente para generar, en el agregado, una curva que relacione el nivel de inversión privada al grado de

desigualdad es suponer que los inversionistas tengan distintas capacidades de absorber pérdidas. Esta relación será positiva si se relaciona inversión con igualdad, en lugar de desigualdad. De una situación inicial fuera de la región de tolerancia E*, la teoría predice que a mayor igualdad habrá mayor nivel de inversión. Cuanto menor el grado de igualdad —y de la correspondiente estabilidad social— mayor será el riesgo de la inversión. Entonces, solo los mayores inversionistas, con la mayor capacidad de absorber las eventuales pérdidas, invertirían. Si el grado de igualdad se incrementa, el riesgo cae y ahora los inversionistas con las menores capacidades de absorber las eventuales pérdidas entrarían a invertir. Es claro que, una vez que el grado de igualdad ingresa a la región E*, la curva se vuelve horizontal. Los cambios en la igualdad ya no tienen efecto sobre la inversión.[3]

Por comparación con los resultados obtenidos aquí, se puede decir que existen tres teorías de la inversión privada en la literatura. La primera sostiene que la inversión es endógena al proceso económico. La segunda argumenta que es endógena tanto al proceso económico como al proceso socio-político. La tercera teoría dice que es totalmente exógena. En el primer caso, la teoría establece que la inversión privada depende de la tasa de interés y el retorno esperado, que, a su vez, es determinado por los precios relativos (Barro 1997). El segundo caso introduce el supuesto de que el retorno esperado también depende del grado de estabilidad en el sistema socio-político (Alesina y Perotti 1996, Figueroa 1993). En el tercer caso, los inversionistas son guiados simplemente por sus «instintos animales», como argumentó Keynes.

La segunda teoría es la adoptada aquí. La inversión depende no solo del proceso económico sino, también, del proceso so-

[3] A grados de igualdad excesivos, la curva se volverá decreciente. La curva que relaciona inversión con igualdad tiene entonces una forma de U invertida, con un segmento plano antes de decrecer.

cio-político de la sociedad, es decir, del grado de orden social. La adopción de la tercera teoría habría implicado que el crecimiento económico dependa del estado de ánimo de los capitalistas y que la inversión fuese exógena y el crecimiento económico un evento, no un proceso. Esta predicción no es consistente con los datos del mundo real, como se mostró arriba.

Con el fin de ilustrar el equilibrio dinámico del producto total (Y) y su distribución (D), en las tres sociedades bajo estudio, los sectores de subsistencia deben ser añadidos al sector capitalista para tener las sociedades omega y sigma. Sin embargo, tanto en omega como en sigma, la acumulación de capital ocurre únicamente en el sector capitalista. Los sectores de subsistencia son residuales. La forma reducida de la función de inversión presentada anteriormente se aplica también a las sociedades omega y sigma.

La teoría de la inversión privada presentada predice, entonces, que, en una sociedad capitalista, el flujo de inversión privada de largo plazo, aquel que genera crecimiento económico, depende de su *stock* de bienes públicos. Pero, a la luz de la teoría propuesta aquí, este factor es solamente la causa próxima. ¿De qué depende el *stock* de bienes públicos? Este *stock* depende de los *stocks* iniciales y de la desigualdad inicial. La causa última en el crecimiento de una sociedad capitalista está en sus condiciones iniciales. En las teorías omega y sigma, la acumulación de capital toma lugar solo en sus sectores capitalistas, pero la inversión y el crecimiento del producto total dependen de su grado de desigualdad global; dicho grado incluye los sectores de subsistencia.

Los valores del equilibrio dinámico del producto y la distribución para los tres tipos de capitalismo pueden representarse por las siguientes ecuaciones:

$$Y^{o}(t) = \Phi(K_{b0}, K^{i}_{b0}, K_{h0}, \delta_0; S_h, S_m, z^*, r^*; t) \qquad (8.7)$$
$$D^{o}(t) = G(\delta_0)$$

La primera ecuación muestra el equilibrio dinámico para el ingreso nacional. Para variables exógenas fijas, el equilibrio dinámico se puede presentar como una curva creciente a lo largo del tiempo. La dinámica comparativa indica que todas las variables exógenas tienen efectos positivos sobre esa trayectoria (desplazan la curva hacia arriba), con la sola excepción de la desigualdad inicial y el tipo de interés internacional que tienen un efecto negativo. La segunda ecuación indica que la desigualdad del ingreso (D) no tendrá tendencia temporal, pues su valor en el estado estacionario está determinado por la desigualdad inicial en la dotación de activos.

En suma, la trayectoria del desarrollo económico de una sociedad capitalista depende, entre otros factores, de sus condiciones iniciales —entre las que se incluye su desigualdad inicial—. Este efecto es negativo y opera a través de dos mecanismos: (a) la exclusión social, que bloquea la formación de capital humano (teoría sigma); (b) el desorden social, que bloquea la inversión privada (teoría de la tolerancia limitada a la desigualdad). Sociedades que nacieron al capitalismo con un alto grado de desigualdad no serán tan desarrolladas como las que nacieron menos desiguales.

8.6. Consistencia empírica: la economía mundial

Las proposiciones-beta presentadas arriba pueden ser contrastadas empíricamente ahora. Sobre la teoría de la tolerancia limitada a la desigualdad, el modelo presentado aquí sería refutado si se observara que países con diferentes grados de desigualdad tienen el mismo grado de orden social, o que los incrementos drásticos en la desigualdad no crean desorden social.

Un estudio reciente del Banco Mundial ha mostrado las relaciones estadísticas entre la desigualdad del ingreso y la violencia. La primera variable se mide por el coeficiente de Gini y se basa en los datos del mismo Banco publicados por Deiniger y

Squire (1996). La segunda variable se mide por la tasa de criminalidad sobre la base de los datos de Naciones Unidas publicados en *World Crime Surveys*. La muestra incluye 45 países (16 del primer mundo) para las regresiones sobre homicidios y 34 países (14 del primer mundo) para las regresiones sobre robos. Ambas regresiones excluyen a los países del África debajo del Sahara debido a la falta de datos. El periodo bajo análisis comprende 1965-1995. La principal conclusión de este estudio es que la desigualdad tiene una robusta y positiva asociación con la incidencia de crímenes y robos (Fajnzylber, Lederman y Loayza 2002).

Otro estudio, el de Bourguignon (2000), también encontró la misma asociación estadística. Él basó su análisis en una muestra de 50 países, usando datos de tasas de crímenes para el periodo 1985-1995 y coeficiente de Gini para 1985. Estos dos hallazgos son consistentes con las tres primeras predicciones de la teoría de la tolerancia limitada a la desigualdad.

Como se ha mostrado en el capítulo previo, Latinoamérica es la región donde la desigualdad es la más alta en el mundo. Esta región presenta ciertas características particulares respecto a las otras regiones. Las políticas del Estado dirigidas a incrementar el ingreso de los pobres han sido discutidas en esta región de manera más intensiva. Esta política ha sido llamada «populismo económico». Como Sachs (1990) ha dicho: «La elevada desigualdad de ingresos en Latinoamérica contribuye a la intensa presión política de usar políticas macroeconómicas para incrementar el ingreso de los grupos de bajos ingresos, lo que a su vez contribuye a malas elecciones de política y a un débil desempeño económico» (1990: 10). El populismo económico es entonces endógeno, una respuesta a la desigualdad que va más allá de la tolerancia social. Pero todo indica que la redistribución vía el Estado no ha sido suficiente para restaurar el orden social.

La década del ochenta ha sido llamada la «década perdida» para las economías latinoamericanas. El PBI real per cápita dis-

minuyó significativamente durante ese periodo. Pero la caída en los salarios reales fue aún más drástica. Los rostros de las ciudades de Latinoamérica también han cambiado drásticamente en las últimas dos décadas: montos enormes de recursos para defender la propiedad son usados por las empresas y las familias. Parece que, en efecto, cuando el Estado no redistribuye el ingreso de manera suficiente, tomará lugar una redistribución privada y violenta. Esta relación entre la drástica caída en los salarios reales, grado de violencia económica y disminución en el nivel del producto es consistente con la tercera proposición beta.

Acerca de la cuarta proposición beta —demasiada desigualdad ocasiona que las sociedades funcionen con inestabilidad política—, Alesina y Perotti (1996), en una muestra de 70 países con observaciones del periodo 1960-1985, encontró una correlación positiva entre la desigualdad y la inestabilidad socio-política. Latinoamérica es nuevamente la región donde el grado de inestabilidad política es más alto. Los 16 países latinoamericanos en la muestra tuvieron los más altos niveles de desigualdad e inestabilidad socio-política.

Sobre la conducta de los inversionistas, la información estadística por países es limitada, de modo que no es posible poner a la prueba empírica la teoría propuesta aquí. Uno de los pocos estudios que se puede mencionar es el de de Alesina y Perotti (1996). Ellos encuentran una asociación estadística significativa que era negativa entre la inestabilidad política y la inversión, hecho que, sumado al hallazgo mencionado arriba, daría como forma reducida una relación negativa entre inversión y desigualdad. La interpretación que ellos realizan de sus resultados y las conclusiones de política que ellos esbozan, sin embargo, tienen una imprecisión lógica. En su modelo, ellos toman la distribución del ingreso como una variable independiente; es decir, ellos suponen que la distribución del ingreso es exógena. Pero, en su análisis estadístico, ellos utilizan la distribución del ingreso del periodo inicial de la muestra —1960—. Y para las

otras variables, ellos utilizan los valores medios del periodo 1960-1985. Entonces, la interpretación que ellos realizan de sus resultados —«la desigualdad del ingreso incrementa la inestabilidad socio-política, lo cual a su vez disminuye la inversión»— (1996: 18) es inconsistente con su resultado empírico.

Si uno toma esos datos como si estuvieran generados por un sistema dinámico, entonces la relación estadística hallada muestra que la desigualdad *inicial* está negativamente asociada con los niveles de inversión. Este resultado es consistente con las predicciones del modelo teórico presentado arriba.

Existe mayor información para estudiar el comportamiento de la inversión directa extranjera. Los estudios empíricos al respecto muestran que el flujo de inversiones se da mayormente entre los países del primer mundo, de allí se originan las inversiones y allí se dirigen (Markusen 2002, Tabla 1.2). El estudio de Markusen encuentra que los *stocks* de capital de infraestructura y de capital humano de las naciones tienen una asociación estadística significativa con la inversión directa extranjera (2002: 212). Este resultado es también consistente con las predicciones del modelo. La desigualdad no es una variable que se incluya en este análisis estadístico.

La participación de Latinoamérica en la inversión extranjera directa es pequeña y ha caído en las últimas décadas. Sea que la medida relativa incluya a todos los países, o solo a los otros países en desarrollo, la participación de Latinoamérica en la inversión extranjera directa ha caído consistentemente desde los años ochenta, tal como se puede ver en el Cuadro 8.1. Este resultado es consistente con el modelo teórico, pues sugiere que la inversión directa extranjera no se dirige a países muy desiguales.

Sobre la relación entre crecimiento y distribución, una anécdota puede servir para ilustrar la abstracción que los economistas usualmente hacen de la desigualdad inicial como un factor importante dentro de las condiciones iniciales en el proceso de crecimiento económico. En 1993, el profesor Robert Lucas planteó un enigma considerando el hecho de que Corea del Sur y

Cuadro 8.1
América Latina: inversión directa extranjera, 1981-1995

	1981-1985	1986-1990	1991-1995
Todos los países (billones US$)	50	155	209
Países en desarrollo (billones US$)	13	25	72
América Latina y el Caribe (billones US$)	6	8	19
América Latina y el Caribe como porcentaje de los países en desarrollo	46.2	32.0	26.4
América Latina y el Caribe como porcentaje de todos los países	12.0	5.2	9.1

Nota: Los datos se refieren a medias anuales. Los datos de inversión incluyen nuevo capital y reinversiones netas.
Fuente: CEPAL (1995: 55, Cuadro II.1) y CEPAL (1997: 9, Cuadro 1).

Filipinas eran similares respeto a todos los mayores agregados económicos en los inicios de los años sesenta —ambos países tenían similares condiciones iniciales—, pero Corea ha experimentado un crecimiento «milagroso» mientras que Filipinas se ha mantenido estancado. Benabou (1996) se encargó de mostrar que las condiciones iniciales eran realmente muy diferentes respecto a la desigualdad del ingreso. Corea tenía menos desigualdad, una variable que Lucas ignoró. Luego, Benabou resumió los hallazgos de 23 estudios econométricos acerca de la relación entre inversión-crecimiento y distribución, cubriendo países de todo el mundo, y concluyó que ellos «entregan un mensaje consistente: la desigualdad inicial es perjudicial para el crecimiento de largo plazo» (1996: 13).

En resumen, las predicciones de la teoría desarrollada en este capítulo tienden a ser consistentes con los resultados de los

estudios empíricos. Así mismo, la teoría de la tolerancia a la desigualdad que le sirve de sustento también muestra consistencia con los datos de la realidad.

8.7. Desigualdad, inversiones y competitividad internacional

En la formulación de la función de producción presentada arriba —en donde la producción depende del grado de desigualdad—, aparece una nueva categoría de intensidad de factor: una industria particular puede ser considerada más (o menos) intensiva en equidad en comparación a otras industrias. Esta intensidad depende de los requerimientos de los inversionistas por orden social; dichos requerimientos, a su vez, dependen del grado de interrelaciones sectoriales que se requiere en la producción de los bienes. Si el bien puede ser producido de manera aislada, en un enclave, esos requerimientos serán menos significativos. En este caso, el orden social no jugará un papel tan importante en el proceso productivo. Pero en procesos más integrados de producción, el orden social será importante.

Los inversionistas conseguirán sus objetivos de altos retornos económicos explotando las ventajas absolutas, las ventajas comparativas y las ventajas competitivas (comercio intra-industrial) de las economías. Estas ventajas dependen de la dotación de recursos naturales y de bienes públicos. El modelo teórico desarrollado aquí explicaría su lógica en estas decisiones. Pero, a través de sus inversiones, los inversionistas estarían también desarrollando estas ventajas para el futuro. Esto genera el efecto dinámico de las inversiones.

Si el proceso productivo de un bien es menos intensivo en igualdad, entonces la inversión será menos sensible al riesgo-país. Aunque exista desorden social, los capitalistas invertirán en la producción de este bien. En esta categoría se encuentran bienes que se producen en enclaves, como minas, campos pe-

troleros, maquiladoras y centros turísticos. Pero si el producto en cuestión es más intensivo en igualdad, entonces, a un elevado grado de desigualdad y de desorden social, los capitalistas no invertirán en la producción de este bien. En esta categoría se encuentran la manufactura y los servicios.

La competitividad internacional de largo plazo de una sociedad depende de la productividad de su sistema productivo. Es con la productividad que se ganan los mercados internacionales. El incremento en esta productividad requiere un mayor nivel de inversión privada. Pero, entonces, la competitividad internacional depende del grado de desigualdad de la sociedad. Manteniendo todo lo demás igual, un país con mayor igualdad atraería más inversión privada, con lo cual aumentará su productividad y ganará una mejor posición en los mercados internacionales. Debido a su alto grado de desigualdad, la sociedad no podrá desarrollar sus ventajas comparativas ni competitivas.

Medida por su participación en el comercio internacional, Latinoamérica ha perdido competitividad desde los años sesenta. El Cuadro 8.2 claramente muestra esta tendencia. En 30 años, la región ha reducido su participación relativa a casi la mitad, de 7.7% en 1960 a 3.7% en 1992. El comercio entre los países desarrollados explica la mayor parte del comercio mundial y su participación ha crecido. Pero el grupo que ha experimentado el mayor incremento en la competitividad es el grupo conocido como los «Tigres Asiáticos» —países con un grado de desigualdad menor que la de América Latina—.

La importancia de la desigualdad en la determinación de la inversión privada y en la competitividad depende del grado de integración del sector exportador con la economía nacional. La inversión privada busca explotar ventajas absolutas, comparativas y competitivas de acuerdo con las condiciones iniciales de un país en su dotación de recursos naturales y de bienes públicos —donde se incluye la desigualdad inicial—. Una sociedad bien dotada con recursos naturales pero con un alto grado de desigualdad inicial puede solo incrementar las inversiones

Cuadro 8.2.
América Latina: participación en el comercio mundial (porcentajes)

Año	América Latina	Países del Primer Mundo	Tigres asiáticos	Otros
1960	7.7	65.9	3.4	23.0
1970	5.5	70 .9	3.0	20.6
1980	5.5	62.6	6.0	25.9
1990	3.9	71.4	10.1	14.6
1992	3.7	71.5	11.5	13.3

Nota: «Tigres asiáticos» incluye Corea del Sur, Taiwán, Singapur, Hong Kong, Indonesia, Malasia y Tailandia.
Fuente: CEPAL (1995: 35, Cuadro I.4).

que explotan sus ventajas absolutas y algunas de sus ventajas comparativas. Pero estas inversiones no modificarán la desigualdad inicial ni van a disminuir el grado de inestabilidad sociopolítica. Entonces, a este país le sería muy difícil escapar de la trampa de la desigualdad.

Para una sociedad muy desigual, una dotación importante de recursos naturales podría convertirse en «un regalo del demonio». Le sería difícil desarrollarse bajo los cimientos de ventajas absolutas. La explotación de las ventajas absolutas de un país no requiere que este país tenga un orden social. Los bienes para las exportaciones pueden ser producidos en un enclave. Mientras más integrado esté el sector exportador con el resto del sistema productivo, más importante será el orden social. Por esta razón, la inversión privada que desarrolla el sector exportador requeriría de un menor grado de desigualdad.

¿Cuáles son los bienes en los cuales se han especializado los países latinoamericanos? En casi toda su historia y hasta fines de los años setenta, la región se ha especializado principalmente en productos primarios —tales como minerales y petróleo—. La biodiversidad le permite producir productos pesqueros y al-

gunos productos agrícolas —como la coca, café, algodón, azúcar, espárragos y lana—, es decir, productos que no se pueden producir en otro lugar del mundo sin incurrir en altos costos. El turismo también está basado en recursos naturales e históricos.

Una combinación de ventajas absolutas y relativas es la base para la especialización. Pero, con la excepción de algunos productos agrícolas, la explotación de los recursos naturales no ha significado que la región exporte bienes que sean intensivos en trabajo, siendo el trabajo su factor más abundante. Entonces, Latinoamérica exporta, más bien, bienes donde la renta económica es alta. Este hecho implica que una parte importante de estas exportaciones no depende de las variaciones de los precios internacionales, ni del tipo de cambio, ni del salario real, al menos en un rango significativo, pero sí de la inversión.

Todos aceptan que la producción agrícola es relativamente intensiva en trabajo, especialmente en trabajo no calificado. Sin embargo, esta exportación ha perdido relativa importancia de una manera drástica, como se puede observar en el Cuadro 8.3.

Cuadro 8.3
América Latina: estructura de exportaciones (porcentajes)

	1962	1970	1980	1990	1992
Agrícolas	52.4	47.1	31.2	28.3	29.8
Metales y Minerales	13.1	18.2	10.4	11.7	9.7
Combustible	29.1	22.9	40.6	26.5	21.6
Manufacturas	5.2	11.5	17.3	32.9	38.5
Otros	0.2	0.3	0.5	0.6	0.4
Total	100.0	100.0	100.0	100.0	100.0

Fuente: CEPAL (1995 : 39, Cuadro I.6).

Aunque es cierto que las exportaciones manufactureras de la región han crecido, los datos del Cuadro 8.3 ocultan diferencias saltantes entre los países. Cerca al 75% de las exportacio-

nes manufactureras en 1993 provinieron solamente de dos países: Brasil y México. Sin estos países, el promedio simple para la región es 18% en lugar del 39% que aparece en el Cuadro 8.3 (CEPAL 1995: 77, Cuadro III.5).

Dentro de las manufacturas es posible distinguir un sector de bienes tradicionales que son relativamente intensivos en trabajo. Tal es el caso de los textiles, vestido y calzado. Aquí la abundancia relativa de trabajo parece jugar un papel importante en la competitividad. La decisión de algunas compañías multinacionales de ubicar maquiladoras para producir estos productos en varios países de la región está precisamente basada en la abundancia de la oferta de trabajo.

No es posible contar con datos para poder realizar un test estadístico acerca de la hipótesis de que la inversión en sectores primarios requiere de menor estabilidad socio-política. Los datos no existen. Existen grandes corporaciones que, debido a su gran capacidad para absorber altos riesgos, realizan inversiones en todo el mundo, aun en países que sufren disturbios sociales. Este caso es claro en las empresas petrolíferas. Pero las grandes corporaciones son los inversionistas más notables en dichos países.[4]

Algunas conclusiones sobre la relación entre desigualdad y competitividad para una sociedad sigma se pueden mencionar ahora. En esta sociedad, el factor de alto riesgo debe inducir la inversión hacia los sectores con ventajas absolutas; dichos sectores pueden ser explotados en enclaves. La inestabilidad so-

[4] Tomemos el caso de Cusiana, Colombia. Este es, tal vez, el campo petrolero más grande descubierto en Latinoamérica en los últimos 20 años. La compañía petrolera extranjera interrumpió su exploración en una ocasión porque las guerrillas invadieron los campos y los destruyeron con dinamitas y granadas. Pero este campo petrolero será explotado a pesar del hecho que Colombia es un país plagado de pobreza, caos político, guerrillas, y narcotráfico. Este caso sugiere que no existen prerrequisitos de estabilidad socio-política para explotar reservas petrolíferas y otros recursos naturales.

cial derivada de la desigualdad inicial desanima la inversión que incrementaría la productividad de los factores y desarrollaría ventajas competitivas y comparativas. Entonces, es probable que su patrón de comercio internacional esté basado mayormente en ventajas absolutas. El comercio no crearía condiciones para reducir el exceso de oferta de trabajo, especialmente el trabajo no calificado o el mínimamente calificado. El comercio internacional no podría ser un mecanismo para incrementar la igualdad en sociedades tipo sigma.

Las relaciones entre comercio y desigualdad irían en otra dirección. El desarrollo de las ventajas comparativas y competitivas requiere que esta economía tenga un grado de desigualdad poco pronunciado. Esta conclusión se deriva a partir de la lógica de los inversionistas. Ellos pondrán sus inversiones en países que tengan niveles relativamente altos de estabilidad socio-política. Y esta estabilidad depende del grado de desigualdad. Contrariamente a lo que se dice usualmente, las exportaciones no son causa del crecimiento económico; las exportaciones son endógenas en el proceso económico. La inversión genera crecimiento y una productividad más alta, y así es como se gana la competitividad. Mientras los productos de exportación sean más intensivos en igualdad, mayor deberá ser el efecto de la igualdad social en las exportaciones netas. De la teoría de la inversión propuesta aquí, se puede derivar entonces la siguiente proposición beta: *la competitividad internacional de un país depende negativamente del grado de su desigualdad inicial.*

Puesta a la prueba empírica, esta teoría muestra un aceptable grado de consistencia con los datos de la economía mundial. Latinoamérica es la región más desigual en el mundo y tiene una pequeña participación relativa en la inversión extranjera directa y, también, tiene una pequeña porción del comercio internacional. La región exporta productos que son poco intensivos en igualdad y permanece desigual. La trampa de la desigualdad es clara.

La productividad relativa entre países —que ha sido definida aquí como competitividad de largo plazo— depende de la distribución de la inversión a través de los países. Estas inversiones dependen de los bienes públicos —entre los que se incluye el capital humano y el orden social—. Dichos bienes públicos, a su vez, dependen del grado de igualdad inicial. Entonces, la competitividad de una sociedad depende de su grado de igualdad inicial en la dotación de activos económicos y sociales. Debido a que la estabilidad socio-política depende del nivel de la organización de la sociedad para mantener una distribución tolerable, es claro que la competitividad no es simplemente una cuestión macroeconómica o sectorial; es, más bien, una cuestión social.

En un mundo de libre movilidad de capitales, como el de ahora, las sociedades compiten entre ellas para atraer inversiones privadas; de este modo, buscan convertirse en sociedades cada vez más competitivas. Entonces, los países del tercer mundo deben competir entre ellos por la asignación de la inversión privada. Esta competencia es anterior a la competencia en los mercados internacionales de mercancías. Pero los países compiten con sus grados de igualdad. El nacer muy desigual es, para una sociedad, una condición inicial muy desfavorable para esa competencia, especialmente cuando la economía mundial está cada vez más globalizada. Esto es lo que este estudio ha mostrado, teórica y empíricamente.

Capítulo 9
Una teoría general del capitalismo

Este estudio ha tenido por objeto comprender el funcionamiento del capitalismo. Tres tipos de sociedades abstractas han sido construidos para tal efecto: épsilon, omega y sigma. Estas sociedades abstractas intentan explicar los aspectos esenciales del primer mundo y los del tercer mundo. Los datos básicos de estas realidades muestran una consistencia aceptable con estas sociedades abstractas.

La pregunta ahora es saber si se puede construir una teoría general, un conjunto de proposiciones alfa que partan de las teorías parciales por tipo de capitalismo y que no sean contradictorias entre sí. Si se lograra construir tal teoría general, tendríamos una explicación del sistema capitalista en su conjunto. Llegaríamos a la unidad del conocimiento sobre la economía capitalista. Este es el objetivo de este capítulo.

9.1. Teoría general de la inclusión-exclusión

Se hace necesario un breve resumen de las características básicas de las tres sociedades capitalistas estudiadas. En la sociedad épsilon, los individuos están dotados inicialmente con cantidades desiguales de activos económicos (capital físico y capital humano), pero los activos sociales están distribuidos equitativamente. Épsilon es una economía socialmente homogénea. Épsilon es también una sociedad que no está sobrepoblada. ¿Cómo funciona esta sociedad? El instrumento de extracción de esfuerzo es la tasa positiva de desempleo. Los trabajadores que incumplen con la disciplina laboral deben sufrir un costo eco-

nómico, convirtiéndose el desempleo en el instrumento disciplinario. Por lo tanto, el desempleo es una necesidad para el funcionamiento del sistema.

La economía épsilon funciona con desigualdad. El ingreso nacional es distribuido a los capitalistas como beneficios y a los trabajadores como salarios. En la pirámide de ingresos, los capitalistas se encuentran en la cima, seguidos por los trabajadores empleados; y, finalmente, se encuentran los trabajadores desempleados que no cuentan con ingresos. Al igual que el desempleo, la desigualdad es una necesidad para el funcionamiento del sistema. Los trabajadores no pueden ser más ricos que los capitalistas, y los trabajadores asalariados no pueden ser más pobres que los desempleados. El equilibrio estático de la economía épsilon implica valores particulares para el producto total, el desempleo y el grado de desigualdad. Un menor grado de desigualdad no es posible, debido a que los trabajadores están sujetos a la exclusión económica, es decir, están excluidos de los mercados de trabajo, del crédito y de seguros. Esta es, entonces, una situación de equilibrio, en la que los valores de equilibrio se repiten periodo tras periodo si las variables exógenas se mantienen fijas. Estos tres mercados —que juegan un papel fundamental en la reproducción de la desigualdad— se llaman *mercados básicos*. Si las variables exógenas cambian, los valores de equilibrio del producto y del empleo cambiarán, no así los de la desigualdad. La economía épsilon opera con desempleo y desigualdad.

La sociedad omega se diferencia de épsilon en que es una economía sobrepoblada. No podría funcionar como una economía épsilon porque el exceso de oferta es mucho mayor que el desempleo necesario para mantener la disciplina laboral. El funcionamiento del capitalismo requerirá de la existencia de un sector de autoempleo para dicha fuerza laboral excedente —denominado *sector de subsistencia*—.

En omega, el instrumento de extracción de esfuerzo será la condición de que la tasa salarial debe ser mayor que el costo de

oportunidad de los trabajadores, que está dada por su ingreso marginal en el sector de subsistencia. Los trabajadores que incumplen con la disciplina laboral serán despedidos y sufrirán un costo económico, que será la diferencia entre estos dos ingresos. Los trabajadores excluidos del mercado laboral pueden elegir entre el autoempleo y el desempleo. Los primeros están subempleados, en el sentido de que trabajan pero generan ingresos por debajo del salario del mercado, es decir, son parte del exceso de oferta laboral. Cabe señalar que el desempleo resultante no es voluntario, ya que es resultado de una elección económica realizada en una situación de segunda opción. Los trabajadores desempleados no perciben ingresos pero su salario esperado es igual que el ingreso que podrían recibir si se autoemplearan en el sector de subsistencia.

En omega, el ingreso nacional tiene tres componentes: beneficios, salarios e ingresos del autoempleo. La asignación de la fuerza laboral incluye el empleo en el mercado laboral, el autoempleo y el desempleo. Esta economía opera con desigualdad en el ingreso. Existen dos niveles en la desigualdad. La primera es la desigualdad entre capitalistas y trabajadores; y la otra es la existente entre trabajadores: los asalariados tienen mayores ingresos que los autoempleados. Así, la desigualdad es una necesidad para el funcionamiento del sistema. Los asalariados no pueden ser más ricos que los capitalistas ni más pobres que los autoempleados. Esta desigualdad se reproduce debido a las mismas inclusiones y exclusiones de los mercados señalados para el caso de la economía épsilon. La economía omega opera con subempleo y desigualdad. El desempleo puede existir, pero no es una necesidad para el funcionamiento del sistema.

A semejanza de omega, sigma es una economía sobrepoblada. En la economía sigma, los individuos participan en el mercado con una desigual dotación inicial en activos económicos y sociales. Sigma es entonces una economía socialmente heterogénea y jerárquica. Esta es su diferencia con la economía omega. Se trata de una sociedad mucho más compleja que las otras, pues

existen en ellas clases sociales y etnias. Aquí, además de la exclusión económica de los mercados básicos, existe exclusión social. Los individuos que están dotados con menores cantidades de activos políticos y culturales no solo son excluidos de los mercados básicos sino, también, de la provisión de *bienes públicos básicos,* tales como educación y salud, y de los sistemas de protección social y de justicia. Los grupos que son socialmente excluidos enfrentan barreras para su movilidad económica y social.

La sociedad sigma opera como la sociedad omega, con la diferencia de que cuenta con un sector de subsistencia adicional, conformado por los que son excluidos socialmente. De esta manera, el ingreso nacional incluye tanto el ingreso generado en este sector de subsistencia como los mencionados para el caso de la economía omega. Los socialmente excluidos constituyen el grupo con el ingreso más bajo en la sociedad. Ellos operan con tecnología tradicional, debido a su bajo capital humano. Por lo tanto, la productividad de este grupo es la más baja del sistema. Los grupos de trabajadores que son excluidos socialmente no desempeñan papel alguno en el funcionamiento del sector capitalista. Constituyen una fuerza laboral prescindible para el sector capitalista, una subclase en la estructura social.

La economía sigma funciona con desigualdad. Esta desigualdad se reproduce debido a los mecanismos de inclusión y exclusión. Una parte de los trabajadores se encuentra incluida en el empleo asalariado pero otra está excluida del mercado laboral debido a la existencia de la sobrepoblación. Los trabajadores que son excluidos socialmente también son excluidos del mercado laboral. Todos los trabajadores están excluidos de los mercados de crédito y seguros. La economía sigma opera con subempleo, trabajo laboral superfluo y desigualdad. El desempleo puede ocurrir, pero no es una necesidad para el funcionamiento del sistema. La desigualdad es necesaria para el funcionamiento del sistema. Pero el grado de desigualdad que resulta del proceso económico es mayor que el necesario y genera desorden social. Sigma es una sociedad inestable.

Estas son las características de las condiciones del equilibrio estático para cada sociedad. Un resumen del funcionamiento de cada una de estas sociedades capitalistas se presenta en el Cuadro 9.1.

Para generar un sistema dinámico se introduce una teoría de la inversión privada que depende del grado de desigualdad, entre otros factores. El equilibrio dinámico es una trayectoria particular de las variables endógenas de la economía, para valores dados de las condiciones iniciales y de las variables exógenas. En cada economía, el equilibrio dinámico muestra un crecimiento del producto en el tiempo, pero el grado de desigualdad es viscoso en el tiempo. Cambios en las variables exógenas desplazarán la trayectoria del producto; pero no tendrán mayor efecto sobre el grado de desigualdad.

En cada sociedad, todo equilibrio, sea estático o dinámico, será con desigualdad. Aún más, dada la desigualdad inicial en los *stocks* de activos económicos y sociales, el funcionamiento del capitalismo dará lugar a una desigualdad en los *flujos,* en la distribución del ingreso nacional. Dada la desigualdad inicial, la desigualdad en los ingresos no se modificará de manera endógena. Así, la desigualdad inicial se convierte en una característica estructural de toda economía capitalista. Las sociedades que nacieron al capitalismo más desiguales que otras continuarán siéndolo. El sistema capitalista opera, en todas sus formas, con desigualdad económica. Este sistema genera progreso económico y desarrollo tecnológico, lo que incrementa el estándar de vida de la población, pero la desigualdad económica es también una característica del capitalismo.

Otro aspecto característico del capitalismo es que el sistema opera al mismo tiempo con mecanismos de integración y de exclusión. La desigualdad económica no es resultado únicamente de la integración del mercado (de la extracción de esfuerzo a los trabajadores), sino que también es el resultado de la exclusión del mercado. Aún más, sociedades en las que exista mayor grado de exclusión mostrarán mayor grado de desigual-

Cuadro 9.1
Características de las economías épsilon, omega y sigma

Sociedad	Sobrepoblación	Heterogeneidad social	Exclusión económica	Exclusión social	Mecanismo de extración del esfuerzo laboral	Distribución del ingreso nacional y de la oferta laboral
ε	No	No	Sí	No	$u \geq u^*$	$Y = P + wD_h$ $S_h = D_h + U$
ω	Sí	No	Sí	No	$w \geq (1+c)\, v'$	$Y = P + wD_h + vL_s$ $S_h = D_h + L_s + U$
σ	Sí	Sí	Sí	Sí	$w \geq (1+c)\, v'$	$Y = P + wD_{hy} + v_y L_{sy} + v_z L_{sz}$ $S_{hy} = D_{hy} + v_y L_{sy} + U_y$ $S_{hz} = v_z L_{sz}$

dad. El efecto de la exclusión sobre la desigualdad en los ingresos es mayor que el efecto de la explotación.

¿Cuál es la realidad particular que cada modelo teórico intenta explicar? La hipótesis empírica básica es que los países del primer mundo operan como si fuesen economías épsilon, mientras que los países socialmente más homogéneos del tercer mundo operan como si fuesen economías omega. Finalmente, aquellos países del tercer mundo que tienen una herencia de dominación colonial operan como si fuesen economías sigma.

Las características más sobresalientes de cada grupo de países son consistentes con las predicciones del modelo teórico correspondiente. En otras palabras, los países del primer mundo operan con tasas de desempleo positivas y desigualdad económica. Las sociedades más homogéneas del tercer mundo operan con desempleo, subempleo y también con desigualdad económica. Los países del tercer mundo que son multiétnicos y tienen un legado de dominación colonial operan con desempleo, subempleo y desigualdad —en donde la incidencia de la pobreza es mayor en ciertos grupos étnicos—. Tal como lo predice la teoría, en estos países, la gente pobre no es solo gente con bajos ingresos, sino que son individuos étnicamente distintos. La pobreza tiene color de piel, lenguaje, sexo, y localización geográfica. Como en el famoso diálogo entre Fitzgerald y Hemingway, se podría decir que, en los dos primeros grupos de países, la única diferencia entre los ricos y pobres es el dinero; mientras que para el último grupo la diferencia va más allá del dinero.

En consecuencia, hasta este punto, no existen razones para no aceptar las tres teorías desarrolladas en este estudio. Cada modelo teórico es una buena aproximación de la realidad de la economía mundial que intenta explicar.

¿Se podría mostrar que estas tres teorías constituyen teorías parciales de una teoría general del capitalismo? Para ello sería necesario que las tres teorías tengan predicciones empíricas que sean consistentes con los hechos básicos de no solamente cada

tipo de país sino, también, para las diferencias entre países del primer y tercer mundo. La unidad del conocimiento requiere que la teoría general englobe a las tres sociedades estudiadas.

Los supuestos comunes pueden ser identificados con facilidad. Esta teoría general puede ser expresada a través del siguiente conjunto de proposiciones alfa:

α_1: *Contexto Institucional.* Reglas: los agentes participan en el proceso económico dotados de activos económicos y sociales; los activos económicos están sujetos a los derechos de propiedad privados; los agentes intercambian bienes sujetos a las normas del mercado, lo que incluye la norma de que los salarios nominales no pueden disminuir. Organizaciones: firmas, familias y el Gobierno.

α_2: *Condiciones iniciales.* Los individuos tienen una dotación desigual de los activos económicos y sociales. Existen dos clases sociales: capitalistas y trabajadores.

α_3: *La racionalidad económica de los agentes.* Las motivaciones de los individuos son las del *homo economicus.* Los capitalistas tienen dos objetivos ordenados jerárquicamente: primero, la preservación de su posición social y, luego, la maximización de sus beneficios. En el mercado de trabajo, los trabajadores buscan maximizar los salarios y minimizar el esfuerzo. La división de las clases en la sociedad genera conflictos sociales, ya que los trabajadores no son socios plenos de los capitalistas y estos deben utilizar mecanismos para extraer esfuerzo de aquellos. Los Gobiernos buscan maximizar el poder político sujeto a las restricciones del presupuesto público.

α_4: *Relaciones en los mercados.* El sistema del mercado opera con mercados walrasianos y no-walrasianos. No existe mercado para los servicios de los bienes de capital, pues los capitalistas prefieren las utilidades en lugar de las rentas.

α_5: *Tolerancia social a la inequidad.* Los individuos tienen un sentido de justicia acerca de su riqueza absoluta y relativa

en la sociedad; es decir, tienen una tolerancia limitada a la desigualdad económica. Si esta alcanza valores que van más allá del umbral de tolerancia, los individuos reaccionarán y desafiarán las reglas del contexto institucional.

Estos supuestos conforman el núcleo de una teoría general del capitalismo. Dentro de dicho núcleo, el capitalismo puede tomar distintas formas. Las tres sociedades capitalistas abstractas que fueron desarrolladas en este libro se convierten en modelos de esta teoría general.

En realidad, estas sociedades difieren únicamente en dos aspectos. El primero se refiere a la dotación agregada de recursos de la sociedad, es decir, si esta se encuentra sobrepoblada o no. La otra diferencia reside en la desigualdad inicial de la dotación de activos, si los activos sociales (políticos y culturales) están distribuidos desigualmente o no; es decir, si la sociedad es socialmente homogénea o no. Estos factores determinan diferentes contextos, y el capitalismo se desempeñará de forma distinta en cada uno de ellos. Este es el resultado teórico básico que se ha encontrado en esta investigación.

¿Cuáles son las proposiciones beta que se derivan de esta teoría general? ¿Son consistentes con los datos de la economía capitalista mundial? A continuación se busca dar respuesta a estas preguntas.

9.2. La desigualdad de las naciones

¿Qué es lo que predice la teoría general en relación con las diferencias en la desigualdad de los ingresos en los tres tipos de capitalismo?

Consideremos tres casos. En primer lugar, supongamos que las tres economías tienen la misma tecnología y *stock* de capital, supuesto que implicaría que tienen la misma curva de demanda de trabajo. Debido a que la cantidad ofrecida de trabajo

en la economía épsilon es menor a la de omega, la que a su vez es menor que la cantidad ofrecida en sigma, entonces, la ratio capital-trabajo disminuye con el tamaño de la dotación de trabajo. Además, supongamos que el salario real es el mismo en las tres sociedades; entonces, el nivel de empleo en el sector capitalista también será el mismo. De este modo, el exceso de oferta de trabajo será mayor en omega y aún mayor en sigma. Por lo tanto, el grado de desigualdad se relaciona inversamente con la ratio capital-trabajo: el grado de desigualdad en la sociedad épsilon es menor que la existente en la sociedad omega, el que a su vez es menor que el grado de desigualdad existente en la sociedad sigma.

En segundo lugar, supongamos que las tres economías tienen la misma curva de demanda de trabajo, pero que el salario real en la economía épsilon es mayor que el de las demás economías. Entonces, en las sociedades omega y sigma el nivel de empleo en el sector capitalista será mayor que en la economía épsilon. Si, además, se supone que la masa de salarios reales es la misma en las tres economías, entonces en cada economía el grado de desigualdad tendrá el orden anteriormente mencionado.

En tercer lugar, supongamos que la economía épsilon opera con un mayor nivel de conocimiento tecnológico y con un mayor *stock* de capital. La ratio capital-trabajo de cada economía tendrá el mismo orden que antes, pero la curva de demanda de trabajo de la economía épsilon se encontrará por encima de las curvas de demanda de trabajo de las otras economías. El salario real y el nivel de empleo en la economía épsilon serán mayores que en el resto. Adicionalmente, supongamos que la ratio beneficios-salarios es el mismo en todas las economías. La conclusión es la misma: el grado de desigualdad en cada economía mantendrá el orden anterior.

Los tres casos muestran que, bajo un amplio rango de circunstancias, existe un orden definitivo en el grado de desigualdad, siendo la economía epsilon la menos desigual y la economía

sigma la más desigual. La desigualdad de la economía omega se ubicaría entre ambas. Este orden de la desigualdad constituye una predicción empírica —una proposición beta— de la teoría general y puede ser puesta a prueba estadística. La hipótesis empírica es que el grado de desigualdad es menor en los países del primer mundo y es mayor en los países del tercer mundo que se caracterizan por ser multiétnicos; mientras que la desigualdad de los países del tercer mundo que son socialmente más homogéneos se ubica en el medio.

La evidencia empírica se presenta en el Cuadro 9.2, en el que se muestra el grado de desigualdad por grupo de países. Estos cálculos utilizan la nueva base de datos del Banco Mundial que es «la más amplia y cuidadosamente construida», a decir de Furman y Stiglitz (1998). Los países industriales son los menos desiguales, mientras que Latinoamérica y África muestran el mayor grado de desigualdad. En general, este resultado se mantiene con cualquiera de las dos medidas de desigualdad calculadas: el coeficiente de Gini y la ratio de las participaciones de los quintiles extremos en la distribución. En consecuencia, estos datos tienden a corroborar la predicción de la teoría general.

Las regiones de Latinoamérica y África son las más desiguales y son, precisamente, las que tienen un mayor grado de herencia colonial que el resto de regiones (Huntington 1996). Usualmente, las economías capitalistas que son socialmente más heterogéneas y jerárquicas provienen de sistemas de dominación colonial. Así mismo, dentro de la región latinoamericana, los países que son socialmente más homogéneos son menos desiguales (Argentina, Uruguay y Costa Rica), tal como se mostró en el Capítulo 6, Cuadro 6.2.

La teoría general predice que la desigualdad en la distribución del *flujo* de ingresos depende de la desigualdad inicial en la dotación de *stocks* de activos económicos y sociales. Para las tres sociedades, los tres modelos teóricos predicen que el grado de desigualdad en el tiempo es una variable viscosa, que

Cuadro 9.2
Desigualdad por regiones del mundo, 1950-1995

Regiones	Número de observaciones	Coef. Gini (media)	Coef. de Pearson (%)	Ratio de las participaciones en el ingreso de los quintiles extremos (media)
América Latina y el Caribe	110	0.50	12.06	16.02
África Sub – Sahara	40	0.45	20.53	11.61
Este Medio y África del Norte	20	0.41	7.53	7.14
Asia del Este	123	0.36	18.10	7.15
Asia del Sur	60	0.34	13.33	5.50
Primer mundo	238	0.33	17.35	6.63

Fuente: Deininger y Squire (1996: 574-577, Tabla 1).

difícilmente cambia, a menos que exista un *shock* refundacional en la sociedad que cambie el legado de la desigualdad inicial.

La nueva base de datos del Banco Mundial sobre la desigualdad por regiones genera estimaciones que son consistentes con esta predicción. Como indica el Cuadro 9.3, no existe mayor variación entre décadas en la media de los coeficientes de Gini para cada región del mundo.

El Cuadro 9.4 presenta información sobre los coeficientes de Gini por países. A partir de la misma base de datos sobre desigualdad del Banco Mundial, se eligió países con 15 o más observaciones para el periodo 1960-1995. Así, esta muestra está constituida por 8 países (se estableció el límite inferior de 15 para que fuese posible tener un número suficientemente grande de observaciones por país y, al mismo tiempo, un número suficiente de países en la muestra). La idea es probar la invariabilidad de los coeficientes de Gini. Para cada país, este cuadro muestra una baja variabilidad del coeficiente de Gini, medido a través del coeficiente de variabilidad de Pearson. En la muestra, estos coeficientes de variabilidad no se encuentran más allá del 5%, con la excepción de Italia (7%) y Canadá (10%).

Un test no-paramétrico para la diferencia estadística en el valor de las medias presentadas en el Cuadro 9.4 muestra que, en efecto, las diferencias no son estadísticamente significativas *dentro* del grupo constituido por Estados Unidos, Italia y Japón, ni en el grupo de Suecia y Canadá. Sin embargo, las diferencias en las medias *entre* los cinco grupos indicados en el cuadro son estadísticamente significativas. De este modo, Brazil es el país más desigual en la muestra, y esta situación ha permanecido igual a través de los años. El Reino Unido es el país menos desigual y, al igual que en el caso anterior, esto se ha mantenido. Taiwán es el siguiente país menos desigual y ha permanecido en esta situación a través de los años. A diferencia de lo que usualmente se argumenta, la distribución en Taiwán no se ha vuelto menos desigual en el proceso de rápido crecimiento que experimentó; comenzó a crecer con un grado bajo de desigualdad.

Cuadro 9.3
Desigualdad por regiones del mundo, por décadas

Regiones	Coeficiente de Gini (media)			
	1960s	1970s	1980s	1990s
América Latina y el Caribe	0.53	0.49	0.50	0.49
África Sub – Sahara	0.50	0.48	0.43	0.47
Este Medio y África del Norte	0.41	0.42	0.40	0.38
Asia del Este	0.37	0.40	0.39	0.38
Asia del Sur	0.36	0.34	0.35	0.32
Primer mundo	0.35	0.35	0.33	0.34

Nota: los coeficientes de Gini son medias simples, sin ponderar. Algunos cambios dentro de regiones pueden ser causados por diferencias en la muestra entre periodos.
Fuente: Deininger y Squire (1996: 584, Tabla 5).

Cuadro 9.4
Coeficientes de Gini por países seleccionados, 1960-1995

País	Número de observaciones	Coeficiente de Gini				
		Media	Mínimo	Máximo	Desviación estándar	Coeficiente de Pearson (%)
Brasil	15	57.28	53.00	61.76	2.66	4.64
EE.UU.	32	35.49	33.50	41.00	1.42	4.00
Italia	15	34.93	32.02	41.00	2.61	7.47
Japón	23	34.82	32.50	37.60	1.35	3.88
Suecia	15	31.63	27.31	33.41	1.49	4.71
Canadá	21	31.18	27.41	32.97	1.72	5.52
Taiwán	26	29.62	27.70	33.60	1.53	5.17
Reino Unido	31	25.98	22.90	32.40	2.61	10.05

Nota: los países han sido seleccionados de acuerdo con sus diferencias estadísticas en valores medios, para lo cual se ha utilizado la prueba no-paramétrica de Friedman.

Fuente: Deininger y Squire (1996: 584, Tabla 5).

El estudio de Li, Squire y Zou (1998) también demostró que el coeficiente de Gini es relativamente estable en cada país. De la muestra de 49 países, tomados de la base de datos del Banco Mundial, tan solo en siete países el coeficiente de Gini mostró tendencias en el tiempo que fueron estadísticamente significativas y también cuantitativamente importantes. En cualquier caso, en ningún país de la muestra se observó una caída en el grado de desigualdad.

Otro estudio que, utilizando fuentes distintas a las del Banco Mundial, también llegó a resultados similares es el de Atkinson (1996). Este autor encontró en una muestra de 17 países de la OECD, para el periodo 1970-1992, que no había tendencias fuertes de cambios en la desigualdad, excepto para el caso de Reino Unido —y solo para la segunda mitad de la década de 1980—.

Todos estos resultados empíricos son consistentes con las predicciones de la teoría general. Y refuerzan los resultados anteriores. Si el grado de desigualdad es más o menos constante en cada país, el orden de la desigualdad entre países debe serlo también. Los países que fueron los más desiguales en los años sesenta continúan siéndolo en los noventa; los países que fueron los menos desiguales a inicios del periodo permanecieron así al final.

La estabilidad de largo plazo en el grado de desigualdad, en la mayoría de regiones del mundo y en la mayoría de países del primer y tercer mundo, parece indicar que la desigualdad es una característica estructural de cada país. Se puede decir que en cuanto a su grado de desigualdad, los países *son* diferentes, no *están* diferentes. El grado inicial de la desigualdad de un país difícilmente cambia. Los países que «nacieron» bastante desiguales han permanecido en dicha situación. No se ha observado cambios importantes en la posición relativa de los países de acuerdo con su grado de desigualdad en la economía mundial, al menos en las últimas cuatro décadas. Este resultado empírico es consistente con la predicción de la teoría general.

9.3. Orden social y desigualdad

La teoría general predice que la excesiva desigualdad genera desorden social. Esta consecuencia se debe al supuesto de que los individuos tienen una tolerancia limitada a la desigualdad. Una sociedad humana está compuesta por un conjunto de individuos y sus interrelaciones. La calidad de vida de un individuo no puede ser considerada independiente de las acciones de los demás. Además, una sociedad humana se encuentra llena de externalidades. Precisamente, la excesiva desigualdad es una de estas. Cabe señalar que se trata de una externalidad negativa. Así, esta excesiva desigualdad constituye la fuente del desorden social.

La sociedad debe pagar el costo del desorden social bajo la forma de mayores costos de transacción, mayores impuestos para financiar el gasto público destinado a combatir la violencia, y mayores gastos privados en bienes y servicios de seguridad. Otro costo que se genera es la caída de la inversión privada debido al incremento del riesgo país que se encuentra asociado al desorden social. No existe un mecanismo automático del mercado que restaure el grado de desigualdad a sus valores tolerables. La desigualdad económica no se autoregula.

La predicción de la teoría general es que a un mayor grado de desigualdad le corresponde un menor grado de orden social. La evidencia empírica es consistente con esta predicción, tal como se mostró en el capítulo 8.

Según la teoría general, la desigualdad en los ingresos es resultado del funcionamiento del mercado. Si la desigualdad cae dentro de la región de la tolerancia social, no genera ineficiencia en la economía. Si la desigualdad es excesiva, en el sentido de que el grado de desigualdad cae fuera de la región de tolerancia social, este resultado constituye una falla del mercado, pues lleva a una pérdida de eficiencia económica en la economía.

¿Qué acciones toman los Gobiernos en una sociedad donde la desigualdad del ingreso cae fuera de la región de tolerancia?

Uno esperaría que en un capitalismo democrático el mandato de la mayoría restaure el equilibrio automáticamente. Sin embargo, los votantes no son los que eligen las políticas que serán tomadas. Como Banfield sugirió hace un tiempo atrás, las políticas son elegidas por los representantes. Estos representantes no son una muestra aleatoria de los votantes, como los doctores no lo son de sus pacientes (Banfield 1958). Los votantes y los representantes, entonces, enfrentan el problema del principal-agente.

Okun (1975) dio una respuesta a la interrogante propuesta. Él señaló que un capitalismo democrático opera con una doble moral: el sistema político pregona la igualdad pero el sistema de mercado genera desigualdad. Para que el sistema sea viable, la sociedad establece algunas reglas que limitan la desigualdad. El Estado establece derechos que se dirigen a cumplir dicho objetivo. Algunos bienes y servicios son retirados del mercado y distribuidos como bienes públicos. La idea es que el dinero no debería comprar todo en la sociedad. Esta concepción forma parte del contrato social. Así, el orden social es conseguido a través de la democracia.

La teoría de Okun parece explicar el caso de los países del primer mundo. De hecho, estos países han desarrollado sistemas de protección social. Los individuos no pasan hambre porque este riesgo está socialmente asegurado, como un derecho. Tomemos, como ejemplo, el seguro de desempleo. Debido a que la economía opera con desempleo, el Estado establece un derecho que lo compense. De esta manera, el desempleo no generará desorden social. Pero el ingreso compensatorio destinado al trabajador desempleado debe ser menor que la tasa salarial —si la compensación fuese completa, el mercado laboral se transformaría en un mercado walrasiano—. De hecho, es así como funciona el seguro de desempleo en el primer mundo. Por consiguiente, esta compensación (parcial) es consistente con la teoría épsilon en donde el desempleo juega un papel importante en asegurar la disciplina laboral.

Sin embargo, en el tercer mundo, la teoría de Okun no se cumple. En este caso, los sistemas de protección social son bastante débiles y los derechos casi no existen. ¿Por qué sucede esto? ¿Cómo se controla el desorden social? El desorden social es —al parecer— controlado a través de métodos autoritarios. Los Gobiernos utilizan la fuerza para mantener bajo control la violencia política y económica, y la violencia colectiva e individual que se genera cuando existe demasiada desigualdad.

¿Por qué el mecanismo democrático no se toma en cuenta? ¿Por qué no se crean derechos económicos? Una respuesta, sugerida por Maslow, es la teoría de la jerarquía de las necesidades humanas (Maslow 1970). En el tercer mundo, donde el ingreso promedio es bajo, un alto nivel de desigualdad implica que un gran segmento de la población se encuentra en la pobreza absoluta. Las masas son demasiado pobres como para participar del proceso político. Los pobres luchan diariamente para sobrevivir, por lo que no pueden luchar al mismo tiempo para alcanzar bienes que satisfacen necesidades menos urgentes, tales como derechos, libertad de expresión, democracia. La gente que tiene el estómago vacío se interesa por los derechos, pero están demasiado restringidos como para dedicarle sus escasos recursos a obtener este bien público. No es un asunto de preferencias sino de restricciones.

Del mismo modo, los pobres no pueden solventar la acción colectiva para demandar derechos. La razón está, no tanto en el problema olsoniano —falla de la acción colectiva debido a las motivaciones individuales egoístas— sino en el hecho de que la gente es tan pobre que no la puede pagar. De la misma forma, no se trata de un problema de preferencias sino de una cuestión de restricciones del ingreso real.

La otra razón por la que la acción colectiva falla se deriva de la teoría sigma. El costo de la acción colectiva para los excluidos —ciudadanos de segunda clase— es mucho mayor que para aquellos pobres que no son excluidos socialmente. Debido a la segregación, el costo de información es mayor para es-

tos individuos. La falla de las acciones colectivas para demandar derechos es explicada por la pobreza absoluta y la exclusión social. En este caso, es probable que el problema olsoniano sea un problema de segundo orden. Por consiguiente, la falla de la acción colectiva explica la ausencia de los derechos en la sociedad por el lado de la demanda.

Por el lado de la oferta, los Gobiernos de los países del tercer mundo, ¿están interesados en proveer derechos? ¿Cuál es la racionalidad de los Gobiernos en este contexto particular? La economía estándar toma el comportamiento del Gobierno como si estuviera determinado exógenamente. En este estudio, en cambio, el supuesto es que los individuos que forman parte del Gobierno actúan motivados por el propio interés —que ahora incluye no solamente el bienestar material sino, también, el consumo de bienes políticos (poder sobre la gente, prestigio social)—. Esta teoría puede ser re-escrita en términos de un modelo más simple: los Gobiernos buscan maximizar sus votos sujetos a la restricción del presupuesto público.

En las sociedades tipo sigma, este objetivo es conseguido a través de la compra de votos. Este hecho implica relaciones políticas clientelistas. Específicamente, los gobernantes de turno utilizan el presupuesto público en la forma de «gasto social» con el objeto de conseguir votos. Usualmente, estos gastos no crean derechos, precisamente porque es más rentable, en términos políticos, hacerlo como gastos discrecionales y clientelísticos.

¿Se comportarán de una forma distinta los gobernantes de la sociedad épsilon? El supuesto de que buscan maximizar sus votos se mantiene, pero la diferencia reside en las restricciones. En la sociedad épsilon, el Gobierno busca maximizar votos sujetos a la restricción presupuestaria y restringidos por las normas que establecen que una parte del presupuesto debe financiar la provisión de los derechos. En esta sociedad, existe menos espacio para la discrecionalidad del gasto público que es dirigido a la compra de votos.

Esta teoría de los Gobiernos predice que, en las sociedades tipo sigma, la lógica de sus gobernantes no es compatible con el objetivo de modificar la desigualdad del mercado de tal forma que esta se ubique en la región de tolerancia social. Así, estas sociedades no tendrán mecanismos automáticos para establecer el orden social. De este modo, las instituciones democráticas en estas sociedades extremadamente desiguales serán frágiles y tenderán a permanecer en ese estado. Bajo el supuesto de que los umbrales de tolerancia son similares en todos los países, una predicción empírica de esta teoría es que las sociedades tipo sigma funcionarán con instituciones democráticas frágiles y bajo menores grados de democracia, en comparación con las sociedades épsilon. El grado de democracia de las sociedades es, en consecuencia, endógeno.

Los estudios empíricos sobre las relaciones entre la democracia y la desigualdad no son abundantes. Un estudio realizado por Muller (1997) encontró, utilizando un análisis estadístico multivariado, una relación negativa y robusta entre los grados de democracia y la desigualdad para una muestra de 55 países capitalistas. En esta muestra, se incluía 16 países del primer mundo y 39 del tercer mundo, para los años 1960 y 1980. Este resultado es consistente con la predicción de la teoría general.

No se puede tener mayor grado de democracia en una sociedad sigma sin alterar su excesivo grado de desigualdad; es decir, sin alterar el legado de la desigualdad inicial. La teoría general también predice que los derechos de propiedad serán frágiles en una sociedad sigma. Los ingresos no contractuales (la corrupción y la ilegalidad) también constituyen una fuente de ingreso para un segmento importante de la población. Es así como operan las sociedades muy desiguales. El grado de democracia, de respeto a los derechos de propiedad y de corrupción de las sociedades son todos endógenos. En la literatura, sin embargo, algunos autores consideran que estos factores son exógenos, es decir, consideran que estos son factores causales de la falta de desarrollo en el tercer mundo.

9.4. El problema dual de la convergencia

¿Qué es lo que la teoría general tiene que decir sobre el problema dual de la convergencia?

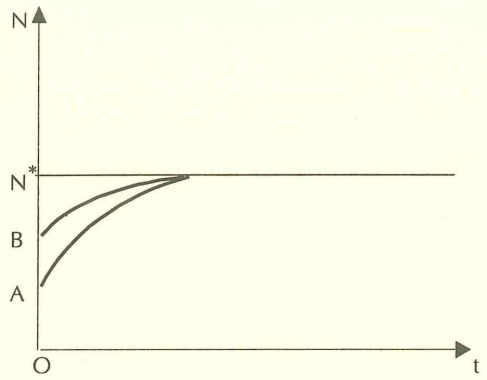
El equilibrio dinámico es, por definición, la trayectoria descrita por la variable endógena del sistema económico en el tiempo, cuando todas las condiciones iniciales y las variables exógenas se mantienen constantes. El movimiento de la variable endógena proviene de su relación intertemporal con otras variables endógenas. El equilibrio dinámico puede ser representado por una trayectoria, es decir, una curva en el tiempo.

Dependiendo de las características del sistema dinámico existen tres tipos de trayectorias. La Figura 9.1 presenta estos tres tipos de equilibrio dinámico para una variable endógena llamada N. En el panel (a), se muestra el caso en el que existe un único valor para el estado estacionario (N^*). Lo que quiere decir que, sin importar el punto de partida, la variable tendrá un único punto de llegada. El panel (b) retrata el caso de equilibrio múltiple (N_1^* y N_2^*), donde el punto de llegada de la variable dependerá del punto de partida. El panel (c) describe el caso en donde no existe un punto de llegada, es decir, el caso en el que, en el equilibrio dinámico, no existe un valor de equilibrio estacionario, donde la trayectoria es una curva creciente. Esto también es un caso de equilibrio dinámico múltiple debido a que la curva de llegada depende de las condiciones iniciales.

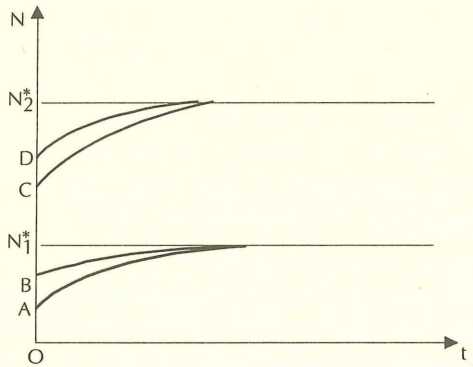
Consideremos, en primer lugar, la convergencia sobre la desigualdad. En cada periodo, el grado de desigualdad es endógeno. Sin embargo, la teoría presentada predice que los cambios en las variables exógenas no afectan mucho la distribución en el ingreso, es decir, que la desigualdad inicial en la distribución del ingreso es reproducida periodo tras periodo. La desigualdad es una variable viscosa. La desigualdad inicial de las naciones difícilmente cambia. Así, las economías del tipo épsilon nacieron relativamente más iguales y continuarán siéndolo. Por otro lado, las economías del tipo sigma nacieron

Figura 9.1
Tipos de equilibrio dinámico

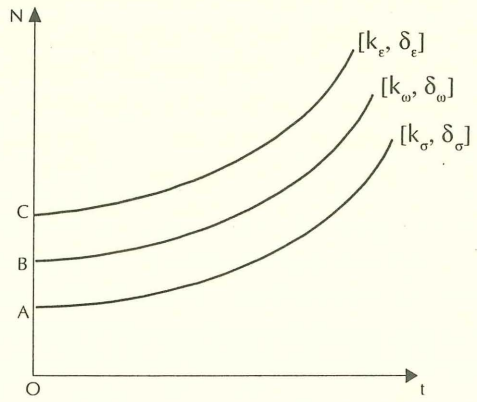
(a)



(b)



(c)



relativamente más desiguales y mantendrán dicha situación; mientras que las economías del tipo omega son un caso intermedio entre la economía épsilon y sigma, y también continuarán con ese grado de desigualdad.

Este sistema dinámico tiene equilibrios múltiples. El panel (b) de la Figura 9.1 representa este caso si el símbolo N es reemplazado por D (el grado de desigualdad en el ingreso). Dado el supuesto de que la sociedad épsilon nació con menos desigualdad que la sociedad sigma ($\delta_\varepsilon < \delta_\sigma$), la curva más baja correspondería a la sociedad épsilon mientras que la curva más alta a la sociedad sigma. La sociedad omega estaría representada por una curva que se encuentre en medio de las dos.

De esta manera, la teoría general predice que no existe convergencia en el grado de desigualdad del ingreso. En todo tipo de sociedad capitalista, el grado de desigualdad en la distribución de los *flujos* de ingreso depende del grado de desigualdad inicial en la distribución de los *stocks* de activos. Esto se representa:

$$D^o(t) = f(\delta_0) \qquad (9.1)$$

Esta predicción es consistente con los datos empíricos. Como se mostró más arriba, los estudios empíricos coinciden en mostrarnos un cuadro donde la desigualdad en el ingreso por regiones y países del mundo capitalista se mantiene casi constante en el tiempo.

Ahora, analicemos el problema de la convergencia del nivel de ingreso. En una economía capitalista, la inversión privada es el motor del crecimiento del producto; ya que esta no solo incrementa el *stock* de capital sino, también, el nivel de tecnología. En la teoría general, la inversión es endógena y depende de la calidad de la sociedad, dada por el *stock* de sus bienes públicos. Estos bienes públicos están constituidos por la infraestructura física, el capital humano y el orden social. La oferta de estos bienes públicos depende del comportamiento del Gobier-

no; dicho comportamiento está determinado por el grado de desigualdad —que es una característica estructural de cada sociedad—. La inversión extranjera tenderá a moverse hacia sociedades de mayor calidad. Estas economías compiten por atraer la inversión con su grado de desigualdad, es decir, con la calidad de su sociedad.

En el Capítulo 8 —ecuación (8.7)—, se mostró que el equilibrio dinámico para el producto total depende del *stock* de capital y la desigualdad iniciales (como condiciones iniciales) y de la oferta laboral, políticas macroeconómicas domésticas y *shocks* externos (como variables exógenas puras). Dados los valores de las condiciones iniciales y de las variables exógenas, el producto total tendrá una trayectoria particular en el tiempo.

La oferta laboral puede ser endogenizada asumiendo que esta depende tanto del nivel del ingreso como de su distribución. Las familias más pobres tendrán más hijos como una estrategia de protección social. Las variables de política macroeconómica pueden ser transformadas en variables endógenas suponiendo que estas son utilizadas como una reacción a los *shocks* externos. Finalmente, los *shocks* externos pueden ser ignorados en la discusión sobre la tendencia de largo plazo del producto.

El equilibrio dinámico del producto dependerá únicamente de las condiciones iniciales, que ahora también incluyen el *stock* inicial de la población. El problema de convergencia en el nivel de ingresos debe ser tratado en términos per cápita (y). Transformando el producto total en producto per cápita, se obtiene el siguiente equilibrio dinámico:

$$y^{o}(t) = f(k_{b0}, k^{i}_{b0}, k_{h0}, \delta_{0}; t) \qquad (9.2)$$

El primer término indica las dotaciones iniciales de capital físico per cápita; el segundo es la dotación inicial de la infraestructura de capital físico per capita; el tercero es la dotación del capital humano per cápita; y el último término es la desigual-

dad inicial en la dotación individual de los activos económicos y sociales. La trayectoria del producto per cápita en el tiempo se encuentra así, para valores dados de las condiciones iniciales, completamente determinada. Es una trayectoria creciente en el tiempo que no converge a un valor estacionario.

En una economía mundial con un mayor grado de globalización, esta función sería única para todas los tipos de sociedades capitalistas. Sociedades que tienen la misma dotación de factores y el mismo grado de desigualdad inicial tendrían el mismo nivel de ingreso con la misma trayectoria a lo largo del tiempo. Si las sociedades difieren en sus condiciones iniciales tendrían diferentes niveles de ingreso y con trayectorias distintas. En particular, sociedades con mayores dotaciones de factores describirían una trayectoria (curva) en el tiempo que tendría una posición relativa superior al de las otras sociedades. Sociedades con un bajo grado de desigualdad inicial mostrarían la misma característica.

Entre las tres sociedades bajo estudio, el equilibrio dinámico de la economía épsilon estará representado por una trayectoria en el tiempo (una curva) que se encuentra por encima de la que le corresponde a la sociedad sigma. En comparación con la sociedad sigma, la economía épsilon empieza con una ratio capital-trabajo, infraestructura de capital y capital humano más altos, y con un menor grado de desigualdad inicial. En este modelo teórico no hay mecanismo que asegure que la curva de sigma converja a la de épsilon. Por el contrario, la inversión per cápita será menor en la economía sigma en comparación con la situación de la economía épsilon. La trayectoria de sigma no conseguirá alcanzar a la de épsilon. La trayectoria del producto per cápita de la economía omega se encontrará en medio de las dos anteriores. Pero esta curva tenderá a converger hacia la trayectoria de la economía épsilon.

El panel (c) de la Figura 9.1 muestra esta predicción de la teoría general, si hacemos que el eje vertical mida el ingreso per capita, y donde únicamente la dotación inicial de los factores y la desigualdad inicial determinan el equilibrio dinámico

del ingreso per cápita. La trayectoria que se inicia en C corresponderá a las economías del tipo épsilon, en B serán las del tipo omega y en A las del tipo sigma. El orden de las curvas está determinado, en parte, por la desigualdad inicial en la distribución individual de los activos (δ_o). Las sociedades que surgieron al capitalismo relativamente más homogéneas tienen un nivel de ingreso mayor y mantienen su posición en el tiempo. El sistema dinámico genera equilibrios múltiples. La curva intermedia tiende a converger a la curva superior, pero no así la curva inferior.

¿Qué es lo que la evidencia empírica muestra acerca de la convergencia en el nivel de ingreso? El estudio realizado por Barro y Sala-i-Martin encuentra «un patrón en la data de corte transversal —para 100 países desde 1960 a 1990— en la que la tasa de crecimiento del PIB per cápita real desde 1960 a 1990 no está relacionada con el nivel del PIB per cápita de 1960» (1995: 7). La tendencia a la convergencia apareció cuando se utilizó información para un conjunto de economías más homogéneas, tales como los estados dentro de EE.UU., las regiones dentro de varios países europeos y las prefecturas en Japón. En estos casos, las zonas más pobres tienden a crecer más rápido que las ricas. Asimismo, se observó que la convergencia también aparecía en el caso de los países de la OECD.

El estudio de Barro y Sala-i-Martin también encontró convergencia *condicional*, es decir, un grupo de países muestra convergencia cuando se controla otras variables relevantes que afectan al crecimiento. Entre estas variables se encuentra el grado de inestabilidad política, que tiene un efecto negativo en las tasas de crecimiento. Este resultado parece ser consistente con lo que predice la teoría general, pues indicaría el caso de convergencia de países hacia una de las curvas de la figura 9.1 —no una convergencia entre estas curvas—, donde el grado de desigualdad inicial subyace al grado de inestabilidad política.

La región del África debajo del Sahara es la que muestra las más bajas tasas de crecimiento en el mundo capitalista. Esta

tragedia ha sido asociada a la diversidad étnica y a la consecuente polarización social de esa región (Easterly y Levine 1997). Si se recuerda que esta región es también una de las más desiguales en el mundo, se comprueba que multietnicidad, inestabilidad política y bajo crecimiento están todos asociados empíricamente con la desigualdad.

Otro estudio empírico introdujo en el modelo de regresiones de Barro, la variable herencia colonial. Sobre la base de una muestra de 77 países del primer y tercer mundo, para el periodo 1960-1988, este estudio encontró que el bajo desempeño de América Latina y de Africa debajo del Sahara en cuanto a crecimiento económico está asociado a su mayor grado de herencia colonial (Bertocchi y Canova 2002).

Cuando se toma una muestra de los países capitalistas del mundo, no se encuentra convergencia en los niveles de ingreso. La convergencia, al parecer, ocurre en el ámbito regional cuando se toma como muestra a las economías del primer mundo. Estas son sociedades relativamente más homogéneas, en las que no existen mecanismos de exclusión social. En contraste, la convergencia entre regiones dentro de los países del tercer mundo parece no ocurrir, aunque la información empírica es escasa.

La teoría general predice una ausencia de convergencia en el grado de desigualdad y en el nivel de ingresos. Los estudios empíricos presentados arriba son consistentes con esta predicción. La evidencia empírica encontrada por Barro y Sala-i-Martin, en la que la inestabilidad política tiene efectos negativos sobre el crecimiento, puede ser puesta en esta perspectiva teórica. La teoría general establece que la inestabilidad política, a través del desorden social, afecta negativamente la inversión privada, pero este factor solamente constituye una causa próxima (una ecuación estructural del sistema teórico). ¿Por qué existe inestabilidad política? La teoría general establece que detrás de la inestabilidad política está la excesiva desigualdad inicial. Luego, la causa última del crecimiento económico hay que encon-

trarla en la desigualdad inicial (una ecuación de la forma reducida del sistema teórico).

Estos resultados teóricos y empíricos indican que el grado de desigualdad es más o menos inelástico en relación con el crecimiento económico. La desigualdad del ingreso por países o regiones es viscosa. No es sorprendente, entonces, que no haya sido posible observar empíricamente las curvas de Kuznets. En las relaciones teóricas entre crecimiento y desigualdad, la teoría general predice que el crecimiento depende de la desigualdad, donde «desigualdad» como variable exógena se refiere a la desigualdad inicial en la dotación individual del *stock* de activos. La tasa de crecimiento de largo plazo depende de la desigualdad inicial que existe en la sociedad. La distribución del ingreso en cada periodo es endógena, por lo que no puede tratársele como un factor causal.

El problema dual de la convergencia, en términos de desigualdad y de nivel de ingreso, tiene dos consecuencias para la economía mundial. La primera es que la distancia económica entre el primer y tercer mundo no se ha alterado en las últimas cuatro décadas. Así, bastaría una sola observación para ver la magnitud de estas diferencias. La Figura 9.2 muestra la posición de las regiones del mundo en términos de desigualdad del ingreso medida a través del coeficiente de Gini y del PIB per cápita. Para el coeficiente de Gini se toma el promedio del periodo 1950-1995, mientras que para el PIB per cápita se toma el año 1998. Es impresionante la distancia que separa a los países industriales del resto de las regiones. Sin embargo, es aún más sorprendente el darse cuenta que esta distancia se ha mantenido casi igual desde los años cincuenta.

En consecuencia, el problema de la convergencia tiene dos dimensiones: los niveles de ingreso del tercer mundo no se incrementan para alcanzar a los del primer mundo, y los grados de desigualdad del ingreso no disminuyen a los niveles alcanzados por el primer mundo. Así como sucede en el espacio newtoniano, existen movimientos en los puntos marcados en la

Figura 9.2 a lo largo del eje horizontal, pero sus trayectorias no tienden a converger.

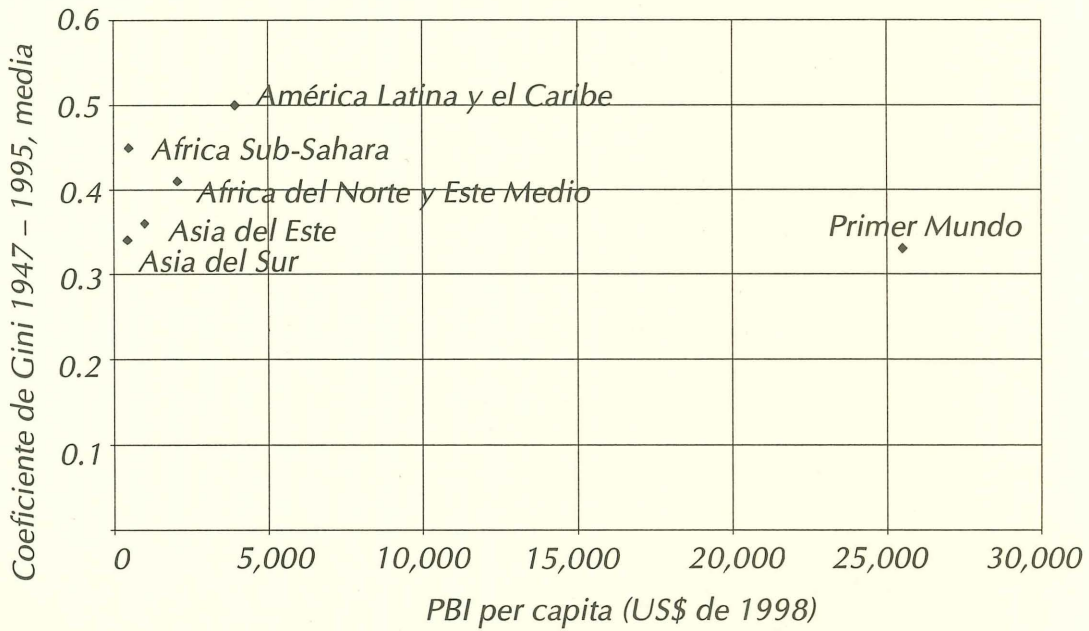
En segundo lugar, la distribución por regiones del ingreso del mundo es muy concentrada y también ha permanecido invariante a través de los años. El primer mundo (Europa occidental, Estados Unidos, Canada, Australia y Japón) representa alrededor del 10% de la población mundial pero concentra cerca del 50% del ingreso mundial (World Bank 2001: Cuadro 1). Este dato es similar al que se encuentra en los países más desiguales del mundo, como Brasil y Perú. Por otra parte, Jones ha mostrado que la desigualdad en el ingreso mundial por países, medido por la ratio del ingreso medio de países ubicados en los quintiles extremos —que es del orden de 25 veces— no se ha modificado en el periodo 1960-1990 (Jones 1998: 64-65).

El politicólogo Samuel Huntington (1996) ha desarrollado una teoría sobre el papel de la cultura y la religión en el desarrollo económico y político de las sociedades. Una predicción de su teoría es que los conflictos religiosos constituyen la mayor amenaza para la paz mundial (1996: 321). La teoría general presentada en este estudio señala otra fuente de amenaza para la paz mundial: la persistencia de la excesiva desigualdad económica entre el primer mundo y el tercer mundo.

Biólogos como Wilson (1998) han sugerido que los recursos del mundo no serían suficientes si el tercer mundo tuviera que alcanzar el nivel de vida del primer mundo; se necesitaría otro planeta Tierra. Visto el orden de magnitud en la desigualdad del ingreso mundial, el nivel de vida del tercer mundo puede elevarse también por medio de la redistribución del ingreso mundial en lugar de hacerlo solo por la vía del crecimiento de la economía mundial. Se puede encarar mejor el problema ecológico con una redistribución selectiva del ingreso mundial.

En suma, el desarrollo económico de las naciones depende, entre otros factores, de sus condiciones iniciales —se incluyen aquí la desigualdad inicial en la dotación individual de los acti-

Figura 9.2
Desigualdad y niveles de ingreso por regiones del mundo



Fuente: Cuadros 9.2 y World Bank (2001, cuadro 1)

vos económicos y sociales—. En el desarrollo económico de las naciones, el peso de la historia es muy importante.

9.5. El desarrollo económico y su evolución

Las sociedades cambian cualitativamente en el proceso de desarrollo económico. Este es el caso de la sociedad omega. En la medida que se produce la acumulación de capital, la economía crece y el exceso laboral tiende a desaparecer. En el muy largo plazo, la economía omega se convertirá en épsilon. En el caso de la economía sigma, el desarrollo económico implicará su transformación, primero, en una economía omega y, luego, en una economía épsilon. La teoría de la exclusión predice que esta evolución se encuentra llena de barreras. Las personas que son excluidas por la sociedad serán dejadas atrás; de modo que el crecimiento del producto no eliminará la sobrepoblación. En las sociedades del tipo sigma, la regla es la no-convergencia, siendo la convergencia la excepción; mientras que en las sociedades omega sucede lo contrario. Pero la mayoría de países del tercer mundo se asemeja a las sociedades sigma.

Así lo indica la experiencia de la economía mundial. Japón es el único caso histórico de convergencia. Este país se convirtió en miembro del club de los países del primer mundo, luego de formar parte del grupo de los países más pobres en 1820 (Maddison 1995). En un futuro cercano, solamente Corea del Sur y Taiwán podrían convertirse en otros ejemplos de convergencia. Cabe resaltar que en estos países existe una baja desigualdad; son una especie de economías omega que se transformarán en economías del tipo épsilon. Bajo el marco de la teoría general, estos casos no son «milagrosos». Debido a que la mayoría de países del tercer mundo se parecen más a la economía sigma, la teoría general predice que no todas las economías capitalistas tenderán de manera endógena a convertirse en sociedades del tipo épsilon.

La trampa que frena el desarrollo de las sociedades sigma reside en las condiciones iniciales de esta sociedad, especialmente en la excesiva desigualdad en la dotación inicial de activos de los individuos. No existe ningún mecanismo en el sistema que pueda cambiar endógenamente el grado inicial de desigualdad. De este modo, la desigualdad se reproduce periodo tras periodo y se convierte en una característica estructural de la sociedad. Bajo el capitalismo, la inversión privada es indispensable para el crecimiento económico. En la sociedad sigma, la inversión no fluirá como lo hace en las sociedades épsilon y omega; además, esta inversión no transformará la sociedad sigma en una omega y luego en una épsilon. El capital humano necesario para adoptar tecnología moderna no puede ser reproducido con facilidad, especialmente por la exclusión social existente. En las sociedades sigma, el potencial para el crecimiento económico no se realizará de forma automática.

Precisamente, en la sociedad sigma es bastante complicado producir capital humano. Como se demostró anteriormente (capítulo 6), en Latinoamérica no existe convergencia en la productividad ni siquiera dentro del sector agrícola, es decir, entre las empresas capitalistas y los campesinos. ¿Qué factor es el que impide que los campesinos adopten la tecnología moderna utilizada por las empresas capitalistas vecinas? Se trata de la ausencia del capital humano. Pero, ¿por qué el capital humano no puede ser acumulado en el campesinado? Esta falla se explica por las exclusiones políticas y culturales existentes.

Como consecuencia de las condiciones iniciales, en la sociedad sigma existen barreras económicas, políticas y culturales para la acumulación de capital humano. Estas barreras hacen que esta acumulación se limite a un segmento de la sociedad solamente. La capacidad de esta sociedad para absorber tecnología moderna es entonces limitada. En general, la transformación de una sociedad sigma en una sociedad omega implica la transformación de trabajadores-z en trabajadores-y. Si esto

no ocurriera, la sociedad sigma tendría una trampa para el desarrollo.

La acumulación de capital humano en la sociedad omega es más viable, ya que se trata de una sociedad más homogénea, en la que no existen exclusiones políticas y culturales. En el largo plazo, los individuos pueden homogeneizarse en términos de capital humano; de modo que la tecnología moderna puede ser adoptada. Así, es posible que las economías del tipo omega puedan aproximarse a la trayectoria del producto de las economías épsilon. Pero, aun así, demasiada desigualdad en los activos económicos bloquea el crecimiento del producto.

Los diferentes equilibrios dinámicos mostrados en la Figura 9.1 implican distintos supuestos acerca del papel que juega la historia. Así, en el panel (a) la historia no es determinante. Sin importar el punto de partida de los países, al final estos tenderán al mismo nivel de producto. Los países que pertenecen al tercer mundo alcanzarían a los países industrializados. En los paneles (b) y (c), la historia sí es importante. El punto de llegada de los países estará determinado por el punto de partida. En los dos casos, existe una trayectoria dependiente de las condiciones iniciales *(path dependence).*

El grado inicial de desigualdad de un país está determinado por un accidente histórico, por un *shock* fundacional. Así, mientras los países del primer mundo nacieron al capitalismo políticamente libres, en la mayoría de los países del tercer mundo el *shock* fundacional fue la independencia de la dominación colonial. Estas condiciones iniciales determinaron diferentes trayectorias de desarrollo económico. Los países que actualmente son países del primer mundo nacieron al capitalismo como sociedades omegas. Tal como la teoría general predice, estos países no enfrentaron barreras para convertirse en países desarrollados. Sin embargo, los países del tercer mundo, en su mayoría, nacieron al capitalismo como sociedades sigma. Existe una tra-

yectoria dependiente en el proceso de desarrollo económico de las naciones; es decir, la historia cuenta.[1]

Este resultado de la teoría general tiene consecuencias sobre la naturaleza de la ciencia económica. Debido a que en el proceso económico la historia cuenta, la Economía es una ciencia histórica. Por lo tanto, la ontología universalista de la Física no se puede aplicar a la Economía.

La conclusión de la teoría general de que la historia cuenta no debe ser tomada como un determinismo histórico. No existe nada en los resultados del proceso económico que sea inevitable. Esta teoría es una teoría social de equilibrio general; no es una teoría física de equilibrio general, como la teoría del todo (*Theory of Everything*) en Física, donde no existen variables exógenas. En efecto, en la teoría general existen variables exógenas que pueden ser utilizadas para cambiar la situación. Esto será discutido en la siguiente sección.

9.6. Las implicaciones de política para las sociedades del tipo sigma

En la ciencia económica, así como en las ciencias sociales en general, las proposiciones de política deben cumplir con algunos requerimientos lógicos. En primer lugar, necesitan basarse en una teoría, pues algunas de las variables exógenas de esa teoría servirán de instrumentos de política. En segundo lugar, la política económica corresponde al campo de las ingenierías en

[1] El biólogo americano Jared Diamond (1999) ha propuesto la hipótesis de que las condiciones iniciales que explican la superioridad del primer mundo se encuentran en su dotación de clima, fauna, flora y geografía. En el proceso de desarrollo de las naciones, el ambiente físico cuenta. La teoría de Diamond permite explicar por qué los países de Europa Occidental fueron los imperios coloniales. Por qué, por ejemplo, los españoles conquistaron a los incas, y no ocurrío al revés. Es una teoría complementaria a la teoría general presentada aquí.

las ciencias naturales. Debe resolver problemas prácticos en realidades concretas, y eso implica un nivel más bajo de abstracción que el de la teoría, en la que el detalle resulta esencial. En tercer lugar, el objetivo que persigue la política no se deriva directamente de la teoría; se establece de manera independiente porque implica utilizar juicios de valor. Finalmente, las proposiciones de política deben resolver la cuestión sobre quiénes serían los actores sociales que llevarían a cabo la acción; para ello, deben indicar la compatibilidad de incentivos. Si no se hace esto, las proposiciones de política serán solo buenos deseos, sin viabilidad política alguna.

Estos requerimientos son tomados en cuenta en las discusiones de política que siguen. Las políticas que se proponen se basan en la teoría general. En un nivel básico, esta teoría ha pasado las pruebas de la falsación. Entonces, se la puede utilizar en el diseño de políticas. La mayoría de los países del tercer mundo se parecen a la sociedad sigma, no a la omega; por consiguiente, se tomará a la sociedad sigma como el referente relevante para discutir las políticas de desarrollo. Pero estas políticas serán muy generales, sin referencia a la situación de países particulares. Se supone que el objetivo de la política es generar mayor nivel de desarrollo económico: mayor nivel de ingreso y menor grado de desigualdad.

La desigualdad inicial en la dotación individual de activos está históricamente determinada. La historia no puede ser cambiada. El objetivo de la política consistiría en cambiar el legado de la historia, es decir, cambiar la desigualdad actual. La herencia colonial tendría que ser eliminada. Esta política implica una redistribución de activos en todas su formas, tanto económicas como sociales. Significaría una transformación cualitativa de la sociedad. Significaría un *shock* refundacional de la sociedad sigma.

¿En qué consiste un *shock* refundacional? Se tiene que hacer una redistribución de los activos económicos, pero, también, una redistribución de los activos sociales. Una forma de trans-

formar una sociedad heterogénea en una homogénea es estableciendo derechos civiles y derechos humanos. Los derechos humanos crean un espacio de igualdad entre las personas. Esta igualdad tendría que ser el requerimiento básico de la organización social, la «igualdad basal» en el concepto de Sen (1995), para que la sociedad pueda desarrollarse sostenidamente. Los derechos económicos implican igualdad en el acceso a los bienes públicos. Los derechos políticos implican que los individuos son iguales ante la ley. Los derechos culturales implican respeto por las diferentes identidades étnicas. Todos estos derechos constituyen los umbrales de la ciudadanía y tendrían que ser establecidos. Esto requerirá de movimientos sociales que luchen por los derechos civiles, como los que se dieron en EE.UU., en los años sesenta, bajo el liderazgo de Martin Luther King.

La literatura económica le ha conferido un gran peso a los derechos de propiedad como prerrequisito para el desarrollo económico. En todo el mundo se reclama el respeto de estos derechos y se busca la mejor forma de protegerlos. Ahora es el momento de reclamar derechos humanos. Como en el caso de los derechos de propiedad, el desarrollo de los derechos humanos puede ser visto como la introducción de una innovación institucional en el sistema —innovación que conducirá a una más rápida tasa de desarrollo económico—.

Para la teoría estándar, el desarrollo económico depende de la inversión privada. Pero, ¿de qué depende la inversión? La respuesta común dada por la economía estándar es que esta depende del entorno económico que, a su vez, depende del comportamiento del Gobierno (Jones 1998, capítulo 7). De este modo, las acciones que tome el Gobierno son consideradas como exógenas al proceso económico. La falla más importante para lograr mayor desarrollo es la falla del Gobierno. Este es el comportamiento que debe ser cambiado por la política económica.

Este paradigma ha sido aplicado en la ola de reformas neoliberales que fueron promovidas por las organizaciones fi-

nancieras internacionales a escala mundial en las dos últimas décadas. Al parecer, este experimento no ha logrado lo que prometió. Esta es la conclusión a la que uno llegaría luego de leer el Informe Anual del Banco Mundial del 2000/2001, titulado *Atacando la Pobreza* (World Bank 2001). La pobreza es aún un gran problema en el tercer mundo, la desigualdad no ha cambiado y el desempeño del crecimiento no ha sido sobresaliente. Este resultado es lo que habría predicho la teoría general. África y Latinoamérica no podrían convertirse en Europa Occidental por más que privaticen las empresas estatales y (re)abran sus economías a los mercados mundiales.

La teoría general se diferencia claramente de la teoría estándar. Si bien la teoría general también supone que la inversión privada depende del ambiente económico, para ella, este ambiente depende del grado de desigualdad de la sociedad. En la teoría general, el comportamiento del Gobierno es endógeno. Debido a la excesiva desigualdad, la sociedad sigma puede tener diferentes tipos de Gobiernos, elegidos y dictatoriales, liberales y estatistas. La inversión podría depender del Gobierno, pero, a la luz de la teoría general, esta es apenas una causa próxima (parte de la ecuaciones estructurales de la teoría); la causa última (la forma reducida del sistema) son las condiciones iniciales, entre las que se incluye la desigualdad inicial.

La teoría general implica que las políticas de desarrollo deberían priorizar aquellas medidas que luchen contra las barreras y contra los mecanismos de exclusión. No se trata de luchar contra la pobreza sino contra los mecanismos subyacentes que la producen y reproducen periodo tras periodo. Este enfoque implica una agenda de política muy distinta. Las reformas neoliberales no consideraron la redistribución de los activos económicos y sociales. Si los pobres fueran ciudadanos de primera categoría, ellos mismos podrían obtener la redistribución de activos económicos. El activo social juega un papel central. Los programas de reforma agraria así lo demuestran. Estos programas redistribuyeron tierras pero no redistribuyeron derechos.

Después de la reforma agraria, la sociedad ha seguido entrampada en la desigualdad y en la exclusión social. El legado del grado *inicial* de desigualdad y la herencia colonial constituyen cargas pesadas para el desarrollo de las naciones pobres.

¿Quién podría llevar a cabo el *shock* refundacional? Aquí, la teoría sigma ha sido presentada utilizando el método del equilibrio general. Sin embargo, la naturaleza del equilibrio es diferente a la que establece la economía estándar. No se trata únicamente de precios y cantidades de equilibrio en los mercados. El concepto de equilibrio es la reproducción social: nadie tiene el poder y el deseo de cambiar la situación.

Los capitalistas ganarían si la sociedad fuese omega en vez de sigma. Sin embargo, no tienen los incentivos para llevar a cabo el *shock* refundacional. Debido a que el orden social es un bien público, ellos no tienen incentivos para producirlo y esperan que otros produzcan un entorno económico más favorable. Los capitalistas se opondrán a cualquier intento de redistribución del capital de su propiedad. Los Gobiernos no tienen incentivos compatibles tampoco. El gasto en derechos no produce mucha rentabilidad política, como sí lo hacen los gastos discrecionales. Al contrario de la economía estándar, la teoría general supone que el comportamiento del Gobierno es endógeno al proceso económico. Así, los capitalistas y los Gobiernos no serán los agentes del cambio. Existe, en ambos casos, un problema de incompatibilidad de incentivos.

Los trabajadores ganarían bastante del *shock* refundacional, pero se encuentran muy restringidos como para llevarlo a cabo. Este es el caso particular de los trabajadores-z. Aquí no existe ningún problema de incompatibilidad de incentivos; más bien, se trata de un problema de restricciones y exclusiones. Los pobres y los excluidos socialmente enfrentan el problema maslowniano de tener que satisfacer primero necesidades primarias, las que no incluyen bienes públicos, como derechos; luego, está el problema olsoniano que limita la acción colectiva; y, por úl-

timo, está el problema de exclusión que significa falta de voz política. El Informe 2002/2001 del Banco Mundial incluye en sus planes de acción la promoción de un mayor poder para los pobres *(empowerment).* Pero la cuestión de cómo lograrlo y con quiénes como agentes de cambio queda sin respuesta.

Es importante notar que en la teoría sigma la dualidad que existe en la economía no se debe a una resistencia de los trabajadores-z por ser incluidos en el sistema; por el contrario, la dualidad es resultado de la exclusión. La cuestión es cómo eliminar los mecanismos de exclusión. La experiencia muestra que la sola redistribución de activos físicos no logra ese objetivo. El fracaso de los programas de reforma agraria en América Latina ilustra este problema. El pensador peruano José Carlos Mariátegui tenía la hipótesis de que «el problema del indio es el problema de la tierra», pero a la luz de la teoría sigma esta idea es muy limitada, pues deja fuera el papel de los activos sociales. En efecto, los programa de reforma agraria que se aplicaron en varios países de América Latina no han creado sociedades omega.

El otorgamiento de algunos derechos solamente tampoco parece producir transformaciones sociales. Esta es la experiencia que muestran los casos de abolición de la esclavitud. En Brasil, Joaquim Nabuco escribió en los años de 1880 que la abolición de la esclavitud tendría que ir acompañada de reforma agraria para que los esclavos ganaran ciudadanía. Aun en Estados Unidos, la liberación de la esclavitud no les dio ciudadanía a los negros y tuvieron que esperar hasta la década de 1960 para lograr ciudadanía con los movimientos por los derechos civiles. En el caso de América Latina es claro que la independencia de España no creó ciudadanía para la población aborigen.

Según la teoría sigma, la transformación social se obtiene redistribuyendo activos tanto económicos como sociales; aún más, para eliminar la exclusión económica se tiene que eliminar, primero, la exclusión social. Si los trabajadores-z tuvieran

derechos de ciudadanía estarían en mejor posición de obtener tierras, crédito, seguros, capital humano de buena calidad. Así se rompería el círculo vicioso exclusión-pobreza. El problema es que ningún agente tendría incentivos para darles su cuota de poder político.

En suma, entre los agentes locales nadie tiene el poder y el deseo para cambiar la solución de equilibrio de bajo nivel que existe en la sociedad sigma. Los individuos, al actuar guiados por el interés personal, conducen a una solución de producción y distribución que es desigual y sub óptima con relación al producto potencial. Esta solución constituye la situación de equilibrio, pero se trata de un equilibrio de bajo nivel. Precisamente, esta es la naturaleza del equilibrio general que hemos desarrollado en la teoría sigma.

Las organizaciones internacionales (organizaciones de Gobiernos del primer mundo y las multilaterales) tienen un poder creciente sobre los Gobiernos del tercer mundo. La ayuda externa que proporcionan a estos Gobiernos constituyen usualmente *shocks* para aumentar el *stock* de capital de infraestructura. Pero no parecen haber logrado ayudar al desarrollo, excepto en algunos países de tipo omega (Taiwán).

El apoyo de las organizaciones multilaterales financieras (especialmente el Fondo Monetario Internacional y el Banco Mundial) a los Gobiernos y las condiciones que exigen constituyen una de las variables exógenas del comportamiento de los Gobiernos. Estas organizaciones podrían ser los agentes del cambio. No obstante, sus preferencias reveladas en sus acciones —especialmente las tomadas en las reformas neoliberales— indican que sus prioridades son otras. Al parecer, en este caso también existe un problema de incompatibilidad de incentivos.

Las organizaciones no gubernamentales siempre son presentadas como una alternativa a las fallas del mercado y del Gobierno. Sin embargo, un estudio sobre el comportamiento de estas organizaciones en América Latina (Meyer 1999) concluye que, a pesar de que las ONGs están interesadas en la pro-

ducción de bienes públicos, estas son manejadas por personas que se guían por sus intereses personales y el oportunismo, «buscando el interés propio con engaño». Debido a que las ONGs dependen de sus donantes, la hipótesis sobre las motivaciones de sus donantes cobra importancia; sin embargo, este tema no ha sido suficientemente investigado.

El estudio de Meyer muestra también que las acciones de los Gobiernos están siendo examinadas más de cerca por las ONGs internacionales, tales como Amnistía Internacional y la Corte de los Derechos Humanos. Como las ONGs son cada vez más internacionales en cuanto a organización, membresía y objetivos, su contribución en la provisión de bienes públicos se encuentra en las actividades que realizan para publicitar el incumplimiento de las normas internacionales por parte de los Gobiernos —normas tales como los derechos humanos y las normas ambientales—.

En suma, se puede fácilmente entender ahora por qué en el tercer mundo la historia de las políticas públicas es la historia de los fracasos para generar desarrollo económico. La falta de desarrollo es un resultado de los fracasos tanto del mercado como de los Gobiernos. Es, también, el resultado del fracaso de las organizaciones internacionales. Nos hemos quedado así sin agentes de cambio que puedan llevar a cabo los *shocks* refundacionales —que constituyen prerrequisitos para un rápido desarrollo económico—. Esta escasez de agentes de cambio —de empresarios schumpeterianos del desarrollo— es el factor limitante fundamental del desarrollo económico.

Apéndice:
Aversión al riesgo frente a aversión a juegos riesgosos

Este apéndice tiene por objeto presentar una teoría sobre la elección económica bajo un contexto de incertidumbre. Para ello, se hace una revisión rápida de las teorías existentes y se fundamenta la necesidad de una nueva teoría. Al final se presenta una teoría que tiene mayor poder de ser sometida a la falsabilidad.

Se define un contexto de incertidumbre cuando los agentes no pueden conocer los resultados de sus acciones. Incertidumbre es falta de información.

Una situación general de incertidumbre puede ser presentada en la forma de una tabla de contingencia, tal como se muestra en el Cuadro A.1. Se supone que los individuos enfrentan *m* eventualidades y *n* activos en los cuales pueden invertir sus recursos financieros dados. Los retornos unitarios en cada evento y en cada activo se representan por el coeficiente a_{ij}. Así, para una asignación de los recursos financieros X en cada activo, el ingreso que se pueden obtener en cada evento también queda determinado. Al parecer el agente conoce todos los datos, tiene toda la información. Entonces, ¿dónde está la incertidumbre? El agente no sabe cuál de los eventos va a ocurrir (siguiendo la definición que introdujo Frank Knight, a esta situación se le llama «un contexto de incertidumbre»).

Supóngase ahora que la distribución de probabilidades de los eventos fuese conocida. En el Cuadro A.1., esta distribución aparece en la segunda columna. Ahora los agentes conocen todos los datos, incluyendo la distribución de probabilidades (a este contexto se le denomina «un contexto de riesgo», siguiendo a Knight).

Cuadro A.1
Tabla de contingencia

Eventualidades	Probabilidades	Activos				Ingreso
		A_1	A_2	...	A_n	
E_1	P_1	$a_{11}X_1$	$a_{12}X_2$	...	$a_{1n}X_n$	Y_1
E_2	P_2	$a_{21}X_1$	$a_{22}X_2$	...	$a_{2n}X_n$	Y_2
...	...	...	...	...	...	...
E_m	P_m	$a_{m1}X_1$	$a_{m2}X_2$	...	$a_{mn}X_n$	Y_m

Dado el contexto mostrado en la tabla de contingencia, el agente enfrenta el problema de elección que se refiere a la asignación de su *stock* dado de recursos financieros a los distintos activos bajo las siguientes restricciones:

$$Y_i = \Sigma_j a_{ij} X_j,\ i = 1, 2, ..., m \quad (1)$$
$$X = \Sigma X_j \quad (2)$$

La cuestión ahora consiste en tener hipótesis teóricas sobre la lógica que guiará al individuo en la asignación de su recurso fijo. Existen varias teoría al respecto y a continuación se hace una revisión de las más importantes.

1. Teoría del ingreso esperado

Esta teoría supone que los agentes buscan maximizar el ingreso esperado. Un modelo simple de esta teoría se puede expresar utilizando las restricciones dadas por las ecuaciones (1) y (2). La condición de equilibrio resultante es la siguiente:

$$\Sigma P_i Y_i = \Sigma_i P_i (\Sigma_j a_{ij} X_j) = \Sigma_j (\Sigma_i P_i a_{ij}) X_j = \Sigma r_j^e X_j \quad (3)$$

El término r^e_j indica el ingreso esperado por unidad del recurso financiero en la actividad *j*. Este coeficiente es igual a la media de todos los posibles retornos en el activo *j*. Ciertamente,

el agente elegirá colocar su inversión en el activo que le produzca el mayor ingreso esperado, es decir, el agente colocará todo su recurso en un solo activo. Y si cambian las restricciones, elegirá otro activo, pero siempre se especializará en un solo activo. Esta proposición beta es empíricamente inconsistente con la observación de que los individuos diversifican sus inversiones.

2. Teoría de Bernoulli

Esta teoría supone que los agentes buscan maximizar el valor esperado de su utilidad. Un modelo simple de esta teoría se puede expresar así:

$$\text{Maximiza } E(U) = \Sigma P_i U(Y_i), \qquad (4)$$

sujeto a las restricciones (1) y (2).

La condición de equilibrio es la siguiente:

$$E(U)_1 = E(U)_2 = \ldots = E(U)_n \qquad (5)$$

El término $E(U)_j = \Sigma_i P_i U'(Y_i)\, a_{ij}$ es la utilidad marginal eperada en el activo *j*. Una unidad adicional del recurso asignado al activo *j* cambiará la utilidad esperada por el aumento en el ingreso y por la contribución de este ingreso al incremento en la utilidad en cada eventualidad. Por lo tanto, el equilibrio estático implica que la utilidad marginal esperada en cada activo debe ser igual a la del resto. Esta teoría predice diversificación en el portafolio de los individuos. Y si algún activo no estuviera incluido en el portafolio, el modelo lo explicaría como un caso en el que toda la curva de la respectiva utilidad marginal esperada está por debajo de todas las otras curvas. Como la utilidad marginal esperada es no-observable, esta respuesta se vuelve irrefutable. No hay manera de saber por qué el individuo elige algunos activos y deja fuera otros.

3. Teoría de la aversión al riesgo

Esta teoría supone que los individuos eligen su portafolio comparando la utilidad esperada del portafolio y la utilidad segura de no jugar; y eligen el valor máximo. Un modelo simple, donde los juegos son equitativos, se puede escribir así:

$$\text{Maximizar: } [U(X), E(U)], \tag{6}$$

$$\text{sujeto a: } E(U) = P_1 U(X+Y) + P_2 U(X-Y)$$
$$P_1 = P_2 = 1/2$$

El equilibrio tiene las siguientes formas:

(a) Si $U(X) > E(U)$, el agente no juega
(b) Si $U(X) < E(U)$, el agente sí juega
(c) Si $U(X) = E(U)$, el agente es indiferente

La solución de equilibrio depende de la forma de la función de utilidad. El caso (a) ocurre cuando la curva de la utilidad es convexa. El valor esperado de la utilidad estará por debajo del valor de la utilidad si el individuo no juega. En este caso se dice que el agente tiene aversión al riesgo, pues prefiere mantener su dotación X con certeza, que mantener esta misma cantidad pero en el promedio, con posibles valores mayores o menores. Debido a la utilidad marginal decreciente en el ingreso, una ganancia en *z* dólares aumenta la utilidad en una magnitud que es más pequeña comparada a la caída en la utilidad cuando pierde *z* dólares. El agente prefiere no jugar.

El resultado opuesto se obtiene si la forma de la función de utilidad es cóncava, es decir, cuando la utilidad marginal es creciente. Este es el caso (b). El agente prefiere mantener la misma cantidad de su recurso pero en promedio, en lugar de mantenerlo con certeza. El agente es afecto al riesgo. Le gusta jugar. Finalmente, el caso (c) corresponde a una función de utilidad lineal. El agente es neutro respecto al riesgo. Le es indiferente jugar o no jugar.

En realidad, estos casos constituyen definiciones de situaciones de equilibrio. Pero hay que recordar que estas definiciones se basan en el supuesto de que los juegos son equitativos, es decir, la probabilidad de los eventos es la misma. Si los juegos no fueran equitativos, el comportamiento de los agentes podrían no corresponder a esas definiciones. El agente que es adverso al riesgo en un juego equitativo puede decidir jugar en un juego sesgado, en el que la probabilidad de ganar sea suficientemente alta. El agente que es afecto al riesgo en un juego equitativo puede decidir no jugar en un juego sesgado, en el que la probabilidad de perder sea suficientemente alta. Que un agente sea o no adverso al riesgo no depende solo de sus preferencias sino, también, de la distribución de probabilidades.

Esta teoría no puede generar proposiciones beta. Este es un caso similar al de la teoría del consumidor, en la que el efecto ingreso no es refutable y solamente sirve para definir si un bien es «normal» o «inferior».

4. Teoría de la media-varianza

Esta teoría supone que el agente tiene una función de utilidad que depende de la media y la varianza de la distribución de probabilidades de los eventos. La media tiene un efecto positivo sobre la utilidad, mientras que la varianza tiene un efecto negativo. La media (M) es entonces un bien y la varianza (V) es un mal. La varianza —o la desviación estándar (S)— mide el riesgo. Esta teoría supone, entonces, que el agente es adverso al riesgo.

Un modelo sencillo de esta teoría se puede expresar como sigue. Supóngase que el agente puede invertir en dos activos. El activo 1 tiene una tasa de retorno fija y segura r_1. La tasa de retorno en el activo tiene una media m_2 y una desviación estándar s_2. La correspondiente tabla de contingencia tiene, en consecuencia, dos juegos, uno con un retorno fijo independiente de los eventos y el otro con la siguiente distribución de retor-

nos: a_{12} en situaciones buenas y a_{22} en situaciones malas. ¿Cómo elige en este contexto el agente? Sea x_1 la participación del activo 1 en el portafolio y x_2 la participación del activo 2, de manera que $x_1 + x_2 = 1$.

El retorno esperado E(r) y la varianza del portafolio se puede mostrar así:

$$
\begin{aligned}
&E(r) = M = r_1 x_1 + E(r_2) x_2 = r_1 x_1 + m_2 x_2, \\
&\text{donde } m_2 = P_1 a_{12} + P_2 a_{22} \\
&V(r) = \Sigma P_i (r_1 x_1 + a_{i2} x_2 - M)^2 \qquad (7) \\
&\quad = \Sigma P_i (r_1 x_1 + a_{i2} x_2 - r_1 x_1 - m_2 x_2)^2 \\
&\quad = x_2^2 \Sigma P_i (a_{i2} - m_2)^2 = (x_2)^2 (s_2)^2 \\
&\Sigma(r) = x_2 s_2
\end{aligned}
$$

Si el agente coloca todo su recurso en el activo 1 obtendrá como retorno r_1 con certeza (con desviación estándar igual a cero). Si asigna todo su recurso al activo 2 obtendrá m_2 en promedio y con una desviación estándar igual a $s_2 > 0$. Si $r_1 > m_2$ el agente colocará todo su recurso al activo 1. Para que el agente invierta algo en el activo 2, este debe tener un mayor retorno esperado.

Estas opciones se muestran en la Figura A.1., que es la forma convencional de representar esta teoría (cf. Varian 1993, Ch. 3). La línea $r_1 L$ es la recta de presupuesto, donde la pendiente indica el precio del riesgo. Si el agente desea una unidad menos de riesgo (en unidades de la desviación estándar), el precio es una menor cantidad de retorno esperado.

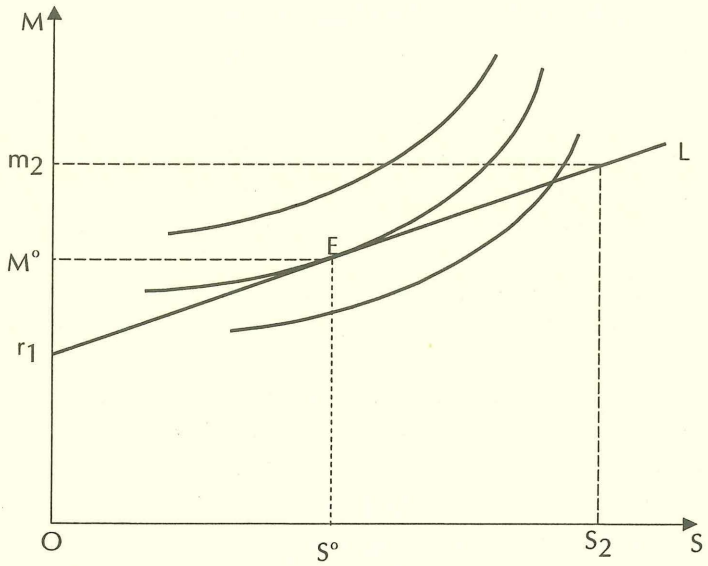
Este modelo se puede expresar de la siguiente manera:

$$\text{Maximiza: } U = U(M, S), \qquad (8)$$

$$
\begin{aligned}
\text{sujeto a: } \quad & M = r_1 + pS \\
& X = X_1 + X_2
\end{aligned}
$$

La condición de equilibrio ocurre donde la recta de presupuesto es igual a la tangente a una curva de utilidad —tal y como ocurre en E, en la Figura A.1—. Se puede mostrar que x_1 es igual a

Figura A.1
Equilibrio en la teoría de la media-varianza



la ratio M_o-r_1/m_2r_1 y x_2 es igual a la ratio m_2-M_o/m_2r_1. La asignación de equilibrio no depende del nivel de recursos X.

Una proposición beta de este modelo es que, si aumenta el *stock* de recursos financieros, el agente los distribuirá en igual proporción a los dos activos. La elasticidad de la demanda por activos es unitaria. Si se observara que la composición del portafolio de los inversionistas depende de su nivel de riqueza, el modelo fallaría empíricamente. La cuestión ahora es si se puede hacer estática comparativa con probabilidades.

5. ¿Se puede utilizar probabilidades como variable exógena?

La definición usual en la teoría estadística es que la probabilidad de un evento E es el valor límite de la frecuencia relativa, es decir, el número de veces que, en una serie de *n* repeticiones

del experimento, tal evento ocurre, cuando *n* tiende a infinito. Esta definición da lugar a dos preguntas: ¿existe el valor límite de la frecuencia relativa?, ¿es este número único? Esta definición supone que ambas respuestas son afirmativas. Pero, más allá de estos supuestos, el hecho es que esta definición no es operativa. La probabilidad no es observable.

Otra definición más empírica de probabilidad se refiere a la frecuencia relativa de los resultados de los eventos. Esta definición también supone que para cada evento E en el espacio de las muestras S existe esta frecuencia relativa y que este número es único. ¿Cómo se calcula la probabilidad? La manera usual se ilustra en los ejemplos que siguen —estos ejemplos han sido extraídos de un texto universitario típico (Ross 1998: 36)—:

(a) Si hacemos un experimento que consiste en lanzar una moneda, y si suponemos que una cara (c) tiene la misma posibilidad de salir que un sello (s), entonces P(c)= P(s)=1/2.
(b) Si la moneda estuviera cargada, y nosotros sintiéramos que una cara tuviera el doble de posibilidades de salir que un sello, entonces P(c)=2/3 y P(s)=1/3.
(c) En otras palabras, si suponemos que todos los resultados de un experimento son igualmente probables, entonces la probabilidad de un evento E es igual a la proporción de puntos en el espacio de la muestra que tiene el evento E.

La segunda definición también hace supuestos sobre el valor de las probabilidades. Pero, ¿cómo se podría hacer la prueba empírica de estos supuestos? ¿Se debería establecer intervalos de confianza para aceptar o rechazar una desviación del valor 0.5, que es el valor de la probabilidad si todos los lados tienen la misma posibilidad de salir? Si es así, se aceptaría que probabilidad es posibilidad. Pero, ¿cuál es la posibilidad de obtener 0.5 en un experimento de lanzar una moneda? Eso depende de la probabilidad. Entonces, posibilidad es probabilidad.

Como señaló Geogerscu-Roegen (1971: 52-59), la probabilidad es un concepto que conduce a un razonamiento circular si

se le quiere colocar en términos operativos; para hacerlo operativo, habría que definirlo de otra manera. Este mismo autor propuso la siguiente condición: la probabilidad de un evento E se puede conocer si, y solo si, se conoce el mecanismo por el cual los eventos son generados. Por lo tanto, la frecuencia relativa de los eventos se podría derivar de las características físicas de los mecanismos que generan los eventos. Si se conoce las características físicas de un moneda y la manera como es lanzada al aire, se puede derivar las probabilidades. La probabilidad de un evento se conocería a priori; aún más, tendría el mismo valor para todos los agentes.

Los propietarios de salones de juego establecen el mecanismo por el cual los eventos serán generados en cada juego; así establecen la distribución de probabilidades. Esta es la manera como las firmas hacen beneficios en la industria de los juegos. Este ejemplo ilustra el principio señalado arriba: el conocimiento de los mecanismos que generan los eventos nos permite determinar la distribución de probabilidades.

En el proceso económico es muy difícil aplicar este criterio. Los mecanismos de generación de los eventos no pueden ser conocidos, pues no son mecánicos. Por ejemplo, para un país, los mecanismos que generan un aumento en la tasa de interés internacional no pueden ser conocidos. En suma, en el proceso económico las probabilidades son no-observables. La distribución de probabilidades no es una *variable exógena,* sino un *elemento exógeno* del proceso. Por lo tanto, la distribución de probabilidades se tiene que suponer como dada, pues, si se la hiciera variar, la teoría se volvería inmune al fracaso en la prueba empírica. Así, la teoría de la media-varianza se vuelve inmortal si se quisiera ponerla a la prueba empírica utilizando probabilidades.

Si la distribución de probabilidades va a ser una variable exógena, tiene que ser observable. O tiene que establecerse un supuesto que haga que esta distribución dependa de una variable observable. Este procedimiento implica introducir otra teoría dentro del sistema teórico inicial.

El proceso económico implica repetición. En este proceso, la frecuencia relativa de los eventos puede ser observada. Por lo tanto, se puede tomar esta definición de probabilidad, como la frecuencia relativa observada. Es esta la probabilidad que se puede utilizar en las teorías analizadas aquí.

6. El punto de desastre de Hicks

Ninguna de las teorías presentadas puede explicar un dato de la realidad: no todos los individuos participan en todos los juegos. Los campesinos juegan a las loterías, pero no juegan a la adopción de cambios tecnológicos.

Un factor que se abstrae en la teoría de la media-varianza es el riesgo de pérdidas significativas para el individuo. En la teoría de la media-varianza, el valor de la desviación estándar puede implicar el riesgo de pérdidas de ingresos absolutos que, de ocurrir, llevarían al individuo al desastre económico.

Hicks (1989) introdujo una nueva teoría de la elección bajo incertidumbre. Su proposición es que los individuos establecen límites a las pérdidas que pueden sufrir en los portafolios que eligen, pues pérdidas muy altas significarían «el desastre» como agentes económicos. Existe entonces un umbral de tolerancia a sufrir pérdidas que el individuo establece. Riesgos de pérdidas que vayan por encima de este umbral serán evitados.

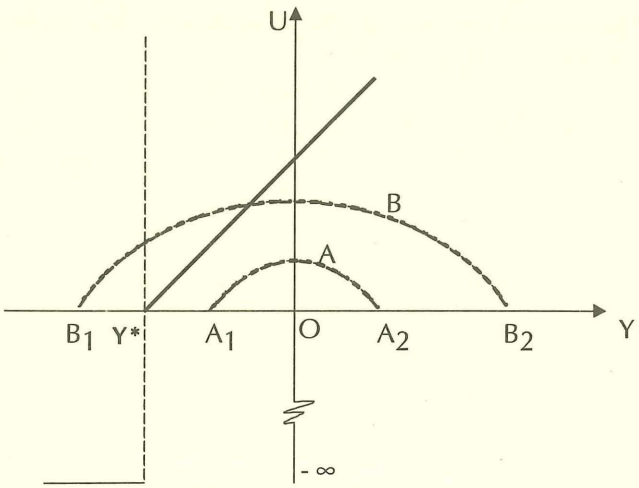
Un modelo de esta teoría se puede presentar así:

$$\begin{aligned} &\text{Maximizar:} && U = f(Y), && (9)\\ &\text{donde:} && U > 0 \text{ si } Y > Y^* \\ & && U = -\infty \text{ si } Y \leq Y^* \end{aligned}$$

La función de utilidad depende del ingreso real Y. La relación es positiva para valores del ingreso superiores al umbral Y^* —umbral que marca el punto de desastre—, pero para valores del ingreso por debajo de este umbral la utilidad da un salto a me-

nos infinito. Este valor de la utilidad es una manera de indicar que el individuo hará todo lo posible por evitar tal situación, pues le significará el desastre como agente económico. La Figura A.2. muestra esta forma de la función de utilidad.

Figura A.2
La teoría de Hicks



La teoría predice que juegos que impliquen pérdidas de ingreso por debajo del umbral no serán jugados por el agente. Esos juegos son muy riesgosos. La racionalidad del agente es la aversión a los juegos riesgosos. Los otros juegos que no impliquen un desastre económico serán jugados. En la Figura A.2, el juego A tiene dos posibles resultados: una pérdida de A_1 y una ganancia de A_2; el juego B muestra esas dos situaciones como B_1 y B_2. El agente jugará el juego A pero no el B. El riesgo de una pérdida en el juego B está por encima de su umbral de tolerancia.

La teoría de Hicks tiene varios problemas. No existe variables exógenas en el sistema. El valor del umbral Y^* es subjetivo. Esta teoría no es falsable. La idea de Hicks parece más útil si el

punto de desastre se introduce dentro de las restricciones de los agentes en lugar de hacerlo en sus funciones de utilidad. Esta modificación se presenta en la sección siguiente.

7. Una teoría lexicográfica de la elección de portafolio

Supóngase que el agente necesita un mínimo de capital X^* para permanecer en el juego, es decir, para seguir siendo el mismo agente económico. Con una cantidad menor, el agente está fuera de juego; por ejemplo, deja de ser capitalista. Este requerimiento está determinado exógenamente. Supóngase, también, que el agente tiene una dotación inicial de capital X, tal que $X>X^*$.

Con estos datos, la pérdida máxima Y^* que el agente podría absorber queda determinada. Esta condición se puede escribir así:

$$Y^* = X - X^*$$
$$Y^* = F(X),\ F'>0 \qquad (10)$$

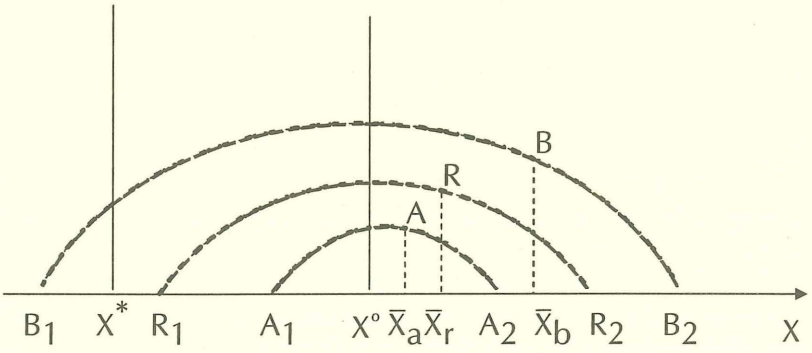
Esta ecuación dice que la capacidad del agente para absorber pérdidas depende de su riqueza. Cuanto mayor su riqueza, el agente puede absorber pérdidas mayores.

La racionalidad del agente es lexicográfica. En primera instancia, el agente decide los juegos que no juega y aquellos que sí juega. El agente deja fuera juegos que implican riesgos de desastre económico, independientemente de sus ganancias posibles. En segunda instancia, en el conjunto de juegos que no implican riesgos de desastre, decide la selección del portafolio. Esta selección se basará en una función objetivo. La primera elección responde a una estrategia de supervivencia del agente: mantenerse en la misma posición de agente; la segunda elección responde a la búsqueda del mayor beneficio.

En la Figura A.3, se presenta esta teoría. Existen dos juegos A y B, con dos eventos en cada caso. El juego R es un portafolio que incluye a ambos. Esta figura es una modificación de la Fi-

gura A.2. —donde se presentó la teoría de la media-varianza—. El valor de la desviación estándar ha sido eliminado y los posibles resultados se han colocado alrededor del valor de la media $\overline{X}$. Luego, el rango de posibles ganancias y posibles pérdidas en cada juego está claramente establecido, rango que no estaba presente en la teoría de la media-varianza.

Figura A.3
Aversión a juegos riesgosos



Está claro que el agente no jugará el juego B solo, a pesar de que este juego le ofrece las mayores ganancias. El agente es muy pobre para jugar este juego. Pero este agente jugará el juego R. El riesgo de pérdida ahora no será B_1, que no podría absorber, sino algo entre B_1 y A_1. Pero, ahora, la ganancia será menor que B_2. La diversificación le reduce el riesgo, pero, a la misma vez, le reduce la media de las posibles ganancias.

En este modelo de la teoría, la variable exógena es la dotación de la riqueza X. El umbral X^* es un elemento exógeno, pues es no-observable. La teoría supone que este punto existe y permanece fijo. La proposición beta del modelo es que a mayores niveles de riqueza, el individuo entrará a jugar nuevos juegos. Existe un orden en los juegos, según el nivel de riqueza de los agentes.

Si se acepta que la selección del portafolio sigue la teoría de la media-varianza, se puede escribir la teoría completa así:

$$\text{Maximizar: } U = U(M,V), \tag{11}$$

sujeto a: $M = f(V)$
$X = \sum X_j$
$X_j = 0$ si existe una pérdida superior a Y_j^*,
tal que $(X - Y_j^*) < X^*$.

La teoría de la media-varianza ha sido transformada así en una teoría completa sobre la elección bajo incertidumbre. Los agentes tienen *aversión al riesgo* pero antes tienen *aversión a los juegos riesgosos*. Por eso, primero eliminan juegos riesgosos del portafolio. Y, dentro del conjunto de juegos que quedan, eligen de acuerdo con la media y la varianza de los portafolios —es aquí donde su aversión al riesgo está presente pero como una elección de segundo orden—.

Con la teoría lexicográfica se puede comprender por qué el campesino compra billetes de lotería de un dólar, pero no compra una vaca genéticamente mejorada de mil dólares. También se puede entender por qué las empresas petroleras son pocas y grandes. Si entran a explorar países en estado de guerra y pierden, su riqueza les permite absorber esas pérdidas y seguir siendo empresas petroleras, pero, si ganan en el juego, habrán añadido a su riqueza inicial un monto muy grande que las vuelve más grandes y, por lo tanto, pueden absorber mayores pérdidas, y así sucesivamente.

De acuerdo con esta teoría, todos los agentes actúan asegurando, en primer lugar, su supervivencia económica. La única diferencia entre ricos y pobres es que los desafíos de supervivencia son distintos: sobrevivir como banquero o como campesino, por ejemplo. Los pobres no pueden ingresar a ciertos juegos que los ricos sí pueden. Por lo tanto, los ricos pueden devenir más ricos. La incertidumbre es un contexto en el cual la desigualdad inicial en activos no podría reducirse endógenamente.

Bibliografía

Alesina, Alberto y Roberto Perotti
1996 «Income Distribution, Political Instability, and Investment». *European Economic Review* 40: 1203-28.

Atkinson, Anthony
1996 «Seeking to Explain the Distribution of Income». En John Hills (ed.), *New Inequalities*. Cambridge University Press.

Banfield, E.C.
1958 *The Moral Bases of a Backward Society*. Nueva York: The Free Press.

Barro, Robert
1997 *Macroeconomics*. 5.ª ed. Cambridge, Mass.: MIT Press.

Barro, Robert y Xavier Sala-i-Martin
1995 *Economic Growth*. McGraw Hill.

Benabou, Roland
1996 «Inequality and Growth». En Bernanke, Ben y Julio Rotemberg (eds.), *NBER Macroeconomics Annual 1996*. Cambridge, Mass.: MIT Press.

Bertocchi, Graziella y Flavio Canova
2002 «Did Colonization Matter for Growth? An Empirical Exploration into the historical causes of Africa's Undervelopment». *European Economic Review*, 46(10): 1851-1871.

Blaug, Mark
1992 *The Methodology of Economics or How Economists Explain*. 2.ª ed. Nueva York: Cambridge University Press.

Blinder, Alan
1987 «Credit Rationing and Effective Supply Failures». *The Economic Journal*, 97, June: 327-352.

Bourguignon, Francois
2000 «Crime, Violence and Inequitable Development». En *Annual World Bank Conference on Development Economics 1999.* Washington DC: The World Bank.

Bowles, Samuel
1985 «The Production Process in a Competitive Economy: Walrasian, Neo-Hobbesian, and Marxian Models». *The American Economic Review*, Vol. 75, N.º 1, March.

Brown, James Robert
2001 *Who Rules in Science?* Cambridge, Mass.: Harvard University Press.

CEPAL
1995 *América Latina y el Caribe: políticas para mejorar la inserción en la economía mundial.* Santiago de Chile: CEPAL.

Deininger, Klaus y Lyn Squire
1996. «A New Data Set Measuring Inequality». *The World Bank Economic Review*, Vol. 10, N.º 3, September.

Diamond, Jared
1999 *Guns, Germs, and Steel. The Fates of Human Societies.* Nueva York: W.W. Norton & Co.

Easterly, William y Ross Levine
1997 «Africa´s Growth Tragedy: Policies and Ethnic Divisons». *The Quarterly Journal of Economics* 107: 1203-1250.

Fajnzylber, Pablo, Daniel Lederman y Norman Loayza
2002 «What Causes Violent Crime». *European Economic Review* 46: 1323-1358.

Feynman, Richard
1985 *Surely, You're Joking Mr. Feynman*. Nueva York: W.W. Norton.

Figueroa, Adolfo
1984 *Capitalist Development and the Peasant Economy in Peru*. Cambridge, UK: Cambridge University Press.

1986. *Productividad y educación en la agricultura campesina de Américal Latina*. Río de Janeiro: Publicaciones ECIEL.

1993 *Crisis distributiva en el Perú*. Lima: Fondo Editorial de la Pontificia Universidad Católica del Perú.

Furman, Jason y Joseph Stiglitz
1998 «Economic Consequences of Income Inequality». En *Income Inequality Issues and Policy Options*. A Symposium Sponsored by the Federal Reserve Bank of Kansas City. Jackson Hole, Wyoming, August 27-29.

Galison, Peter
1997 *Image and Logic. A Material Culture of Microphysics*. Chicago: The University of Chicago Press.

Garraty, John
1978 *Unemployment in History*. Nueva York: Harper Colophon Books.

Georgescu-Roegen, Nicholas
1967 *Analytical Economics*. Cambridge, Mass.: Harvard University Press.

1971 *The Entropy Law and the Economic Process*. Cambridge, Mass.: Harvard University Press.

Glazer, Nathan
1997 *We Are All Multiculturalists Now.* Cambridge, Mass.: Harvard University Press.

Hawking, Stephen
1988 *A Brief History of Time.* Nueva York: Bantam Books.

2002 *The Theory of Everything. The Origin and Fate of the Universe.* Beverly Hills: New Millennium Press.

Hicks, John
1985 *Methods of Dynamic Economics.* Nueva York: Oxford University Press.

1989 *A Market Theory of Money.* Oxford, UK: Clarendon Press.

Hirschman, Albert
1973 «The Changing Tolerance for Income Inequality in the Course of Economic Development». *Quarterly Journal of Economics,* Vol. 87, N.º 4: 544-566.

Huntington, Samuel
1996 *The Clashes of Civilizations and the Remaking of the World Order.* Nueva York: Simon & Schuster.

Jones, Charles
1998 *Introduction to Economic Growth.* Nueva York: Norton.

Kaldor, Nicholas
1957 «Alternative Theories of Distribution». *Review of Economic Studies* 23: 83-100.

Kalecki, Michal
1943 «Political Aspects of Full Employment». Publicado primero en *Political Quarterly* y reproducido después en *Selected Essays on the Dynamics of the Capitalist Economy 1933-1970.* Cambridge, Mass.: Cambridge University Press, 1971.

Keynes, John
1936 *The General Theory of Employment, Interest and Money*. Londres: Macmillan.

Kuhn, Tomas
1970 *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Kuznets, Simon
1955 «Economic Growth and Income Inequality». *American Economic Review* 45: 1-28.

Lewis, Arthur
1954 «Economic Development with Unlimited Supplies of Labor». *The Manchester School of Economic and Social Studies* 28 (May): 139-191.

Li, Hongyi, Lyn Squire, Heng-fu Zou
1998 «Explaining International and Intertemporal Variations in Income Inequality». *Economic Journal* 108(1): 26-43, January.

Lucas, Robert
1981 *Studies in Business Cycle Theory*. Cambridge, Mass.: MIT Press.

1988 «On the Mechanics of Economic Development». *Journal of Monetary Economics* 22 (July): 3-42.

1990 «Why Doesn't Capital Flow from Rich to Poor Countries». *American Economic Review* 80 (2): 92-96.

1992 «On Efficiency and Distribution». *The Economic Journal* 102 (March): 233-247.

Maddison, Angus
1995 *Monitoring the World Economy, 1820-1992*. París: Development Centre, OECD.

Markusen, James
2002 *Multinational Firms and the Theory of International Trade.* Cambridge, Mass.: MIT Press

Marx, Karl
1867 *Das Kapital.* Traducido al inglés como *Capital,* varias ediciones y editoriales.

Maslow, Abraham
1970 *Motivation and Personality.* 2.ª ed. Nueva York: Harper & Row Publishers.

Mayr, Ernest
1991 *One Long Argument. Charles Darwin and the Genesis of Modern Evolutionary Thought.* Cambridge, Mass.: Harvard University Press.

Meyer, Carrie
1999 *The Economics and Politics of NGOs in Latin America.* Londres: Praeger.

Muller, Edward
1997 «Economic Determinants of Democracy». En M. Midlarsky (ed). *Inequality, Democracy, and Economic Development.* Cambridge, Mass.: Cambridge University Press.

North, Douglas
1990 *Institutions, Institucional Change and Economic Per-for-mance.* Cambridge, Mass.: Cambridge University Press.

Okun, Arthur
1975 *Equality and Efficiency. The Big Trade Off.* Washington, DC: The Brookings Institution.

Olson, Mancur
1965 *The Logic of Collective Action. Public Goods and the Theory of Groups.* Cambridge, Mass.: Harvard University Press.

Popper, Karl
1959 *The Logic of Scientific Discovery.* Londres: Hutchinson.

1993 «Evolutionary Epistemology». En M. Goodman y R. Snyder (ed.), *Contemporary Readings in Epistemology.* Nueva York: Prentice Hall.

Prigogine, Ilya
1997 *The End of Certainty. Time, Chaos, and the New Laws of Nature.* Nueva York: The Free Press.

Psacharopoulos, George y Harry Patrinos
1994 *Indigenous People and Poverty in Latin America. An Empirical Analysis.* Washington DC: Aldershot

Rand, John y Finn Tarp
2002 «Business Cycles in Developing Countries. Are they Different?, *World Development.* 30(12): 2071-2088.

Rawls, John
1971 *A Theory of Justice.* Cambridge, Mass.: Harvard University Press.

Ricardo, David
1821 *Principles of Political Economy and Taxation.* Varias ediciones.

Roemer, John
1982 *A General Theory of Exploitation and Class.* Cambridge, Mass.: Harvard University Press.

Romer, David
1996 *Advanced Macroeconomics.* Nueva York: MacGraw Hill.

Ros, Jaime
2000 *Development Theory and the Economics of Growth.* Ann Arbor: The University of Michigan Press.

Ross, Sheldon
1998 *A First Course in Probability*. Nueva York: Prentice Hall.

Sachs, Jeffrey
1990 *Social Conflict and Populist Policies in Latin America*. San Francisco, California: International Center for Economic Growth.

Samuelson, Paul
1947 *Foundations of Economic Analysis*. Nueva York: Atheneum.

Shapiro, Carl y Joseph Stiglitz
1984 «Equilibrium Unemployment as a Worker Discipline Device». *The American Economic Review*, Vol. 74, N.º 3, June.

Schumpeter, Joseph
1934 *The Theory of Economic Development*. Cambridge, Mass.: Harvard University Press. Publicado originalmente en alemán, 1911.

Sen, Amartya
1995 *Inequality Reexamined*. Cambridge, Mass.: Harvard University Press.

Smith, Adam
1776 *La riqueza de las naciones*. Varias editoriales

Stiglitz, Joseph
2002 *Globalization and Its Discontents*. Nueva York: Norton

Stiglitz, Joseph y Andrew Weiss
1981 «Credit Rationing in Markets with Imperfect Information». *American Economic Review*, Vol. 71, N.º 3, June: 393-410.

Thorp, Rosemary
1998 *Progress, Poverty and Exclusion*. Washington, D.C.: Inter American Development Bank.

Varian, Hal
1989 «What Use is Economic Theory». Trabajo presentado a la conferencia *Is Economics Becoming a Hard Science?* Octubre 29-30, 1992. París. El artículo está disponible en la siguiente dirección electrónica: <www/sims/berkeley/edu>

1993 *Intermediate Microeconomics.* Nueva York: N.W. Norton & Company.

Walras, Leon
1954 *Elements of Pure Economics.* New York: A. Kelly Publishers. Publicado originalmente en francés, 1883.

Webb, Richard
1977 *Government Policy and the Distribution of Income In Peru.* Cambridge, Mass.: Harvard University Press.

Weisskoff, Richard y Adolfo Figueroa
1976 «A Comparative Review of Income Distribution in Latin America». *Latin American Research Review,* Vol. XI, N.º 2, pp. 71-112.

Wilson, Edward
1998 *Consilience. The Unity of Knowledge.* Nueva York: Alfred A. Knopf.

World Bank
2001 *World Development Report 2000/2001. Attacking Poverty.* Oxford University Press.

Se terminó de imprimir en los talleres gráficos de
Tarea Asociación Gráfica Educativa
Pasaje María Auxiliadora 156 - Breña
Correo e.: tareagrafica@terra.com.pe
Teléf. 424-8104 / 332-3229 Fax: 424-1582
Octubre 2003 Lima - Perú

Próximas coediciones del Fondo Editorial de la PUCP y el Fondo de Cultura Económica

---

La constante macabra o cómo hemos desalentado a generaciones de alumnos
André Antibí

Los orígenes del gobierno representativo en el Perú. Las elecciones (1809-1826)
Valentín Paniagua

Nueve miradas sin dueño. Ensayos sobre la modernidad y sus representaciones en la poesía hispanomericana y española
Eduardo Chirinos

La sociedad sigma es una sociedad hipotética, en la que los individuos participan en el proceso económico dotados no solo de activos económicos (como supone la economía estándar) sino, también, de activos políticos y culturales. Esta dotación inicial de activos está desigualmente distribuida entre los individuos. Sigma es una sociedad capitalista pero socialmente heterogénea y jerárquica. ¿Cómo funciona esta sociedad en cuanto a la producción y distribución de los bienes? Figueroa elabora un modelo teórico para responder a esta pregunta y muestra, además, que la economía estándar supone, por contraste, sociedades socialmente homogéneas, cuyo funcionamiento obedece a otras reglas de producción y distribución. Debido a que la mayoría de los países del tercer mundo se parecen a la sociedad sigma, esta es una teoría relevante para entender el problema del desarrollo económico. Con ella se puede entender las razones que subyacen a las disparidades persistentes que se observan entre el primer mundo y el tercer mundo. Las políticas de desarrollo que se derivan de la teoría sigma son muy diferentes a aquellas que se derivan de la economía estándar y que son las que se han venido aplicando en el tercer mundo con resultados notoriamente insatisfactorios.

***

«¿Por qué el desarrollo económico no ocurre de manera natural como predice la teoría económica? La famosa teoría de la dualidad de Arthur Lewis supuso que el desarrollo era solo un problema de modernización, en el que lo esencial era la superación de la restricción tecnológica antes que la de la restricción socio-política. La teoría sigma de Adolfo Figueroa explica, por contraste, cómo es que las desigualdades iniciales conducen a un equilibrio de bajo nivel, a una trampa de subdesarrollo en los países del tercer mundo. Es una teoría del desarrollo que combina factores económicos y sociales en un enfoque holístico. [...] Representa una novedosa ruptura teórica no solo con los modelos de Lewis sino con los de toda la economía estándar».

James Copestake. *Theorising the Links Between Social and Economic Development: The Sigma Economy Model of Adolfo Figueroa* (University of Bath, Inglaterra).

ISBN 9972-42-597-5