Lecciones de economía general e internacional

Lecciones de economía, vol. 1 · Pensamiento Económico

Arturo Lorenzo

Edición web 1.0 · noviembre de 2026

Tabla de contenidos

Nota **6**

A1. Objeto y método de la ciencia económica **7**
Introducción . . . 7
1. Objeto de la ciencia económica . . . 9
1.1. Concepto de objeto de una disciplina . . . 9
1.2. Objeto de la economía . . . 9
1.3. Conclusiones . . . 13
2. Método de la ciencia económica . . . 14
2.1. Método científico . . . 14
2.2. Método económico . . . 21
2.3. Conclusiones . . . 27
3. Cuestiones y debates actuales . . . 28
3.1. Economía experimental (1960-actualidad) . . . 28
3.2. Economía conductual (1970-actualidad) . . . 30
3.3. Economía de la complejidad (1987-actualidad) . . . 34
3.4. El método de la macroeconomía (1976-actualidad) . . . 36
Conclusiones . . . 39
Anexos . . . 40
A.1. Corrientes posteriores de la filosofía de la ciencia . . . 40
A.2. Metodología de la escuela austriaca . . . 40
A.3. Preferencias sociales (Güth, Schmittberger y Schwarze, 1982; Fehr y Schmidt, 1999) . . . 40

A2. Los economistas clásicos y Marx **41**
Introducción . . . 41
1. La formación y maduración del pensamiento clásico: A. Smith, T. R. Malthus, D. Ricardo (s. XVIII a s. XIX) . . . 43
1.1. Adam Smith (*La riqueza de las naciones*, 1776) . . . 43
1.2. Thomas Robert Malthus (*Ensayo sobre la población*, 1798) . . . 49
1.3. David Ricardo (*Principios de economía política y tributación*, 1817) . . . 52
1.4. Conclusiones . . . 57
2. La transición al pensamiento neoclásico: el pensamiento clásico tardío (s. XIX) . . . 58
2.1. Nassau Senior (*Esbozo de la ciencia de la economía política*, 1836) . . . 58
2.2. John Stuart Mill (*Principios de economía política*, 1848) . . . 58
2.3. Conclusiones . . . 62
3. Las críticas al pensamiento clásico: Karl Marx (*El capital*, 1867) . . . 63
3.1. Rasgos generales y teoría de la historia: el materialismo histórico . . . 63
3.2. Teoría del valor . . . 65
3.3. Teoría de la distribución de la renta: plusvalía y explotación . . . 66
3.4. Teoría del crecimiento y de los ciclos: las leyes del movimiento capitalista . . . 66
3.5. Implicaciones de política económica: el *Manifiesto comunista* (1848) . . . 69
3.6. El fracaso de las economías comunistas y de las teorías del valor-trabajo . . . 69
3.7. Conclusiones . . . 71

Conclusiones . . . 72
Anexo. Antecedentes: mercantilismo y fisiocracia (ss. XVI-XVIII) . . . 73
A.1. Mercantilismo . . . 73
A.2. Fisiocracia . . . 73

A3. Los economistas neoclásicos **75**
Introducción . . . 75
1. Esfera austro-alemana . . . 77
1.1. Antecedentes: proto-marginalistas en la esfera austro-alemana . . . 77
1.2. Carl Menger (*Principios de economía*, 1871) . . . 78
1.3. Sistematización: Wieser y Böhm-Bawerk . . . 79
1.4. Evolución posterior: Mises y Hayek . . . 81
1.5. Conclusiones . . . 82
2. Esfera anglosajona . . . 84
2.1. Antecedentes: el proto-marginalismo británico . . . 84
2.2. William Stanley Jevons (*The Theory of Political Economy*, 1871) . . . 84
2.3. Sistematización: Alfred Marshall (*Principios de economía*, 1890) . . . 86
2.4. Evolución posterior: Edgeworth, Pigou, la escuela de Cambridge y la revolución ordinal . . . 95
2.5. Evolución posterior en Estados Unidos: Clark, Fisher y Knight . . . 96
2.6. Conclusiones . . . 100
3. Esfera franco-italiana . . . 101
3.1. Antecedentes: el proto-marginalismo francés . . . 101
3.2. Léon Walras (*Elementos de economía política pura*, 1874) . . . 101
3.3. Sistematización: el sistema paretiano (Pareto, 1906) . . . 106
3.4. Evolución posterior: difusión y formalización . . . 106
3.5. Conclusiones . . . 107
Conclusiones . . . 108
Anexos . . . 109
A.1. Diferencias entre la escuela clásica y la neoclásica . . . 109
A.2. Críticas al núcleo teórico: institucionalismo americano y keynesianismo . . . 109
A.3. La escuela de Estocolmo (Wicksell, 1898; Lindahl, Myrdal, Ohlin, Lundberg) 111

A4. Keynes: análisis comparado **112**
Introducción . . . 112
1. Macromodelo neoclásico: una reconstrucción (Cassel, 1918b; Marshall, 1923; Fisher, 1930; Hicks, 1932; Pigou, 1933) . . . 114
1.1. Principales rasgos . . . 114
1.2. Mercado de trabajo . . . 115
1.3. Mercado de bienes y servicios . . . 116
1.4. Mercado de dinero: la teoría cuantitativa (Fisher y Cambridge) . . . 119
1.5. Macromodelo neoclásico OA-DA . . . 120
1.6. Conclusiones . . . 122
2. Pensamiento de Keynes: macromodelo keynesiano (Keynes, 1936; Hicks, 1937; Modigliani, 1944; Hansen, 1953) . . . 123
2.1. Principales rasgos: la sustitución de paradigmas . . . 123
2.2. Mercado de alquiler de trabajo . . . 124
2.3. Mercado de bienes, servicios y dinero: modelo IS-LM . . . 127
2.4. Macromodelo keynesiano OA-DA . . . 134
2.5. Conclusiones . . . 137
3. Reacciones y desarrollos posteriores . . . 138
3.1. El debate sobre los saldos reales . . . 138

3.2. La síntesis neoclásica . . . 140
3.3. Poskeynesianismo . . . 140
3.4. Keynesianos del desequilibrio . . . 142
Conclusiones . . . 149
Anexo. Programas del macromodelo neoclásico . . . 150
A.1. Oferta de trabajo . . . 150
A.2. Demanda de trabajo . . . 151

A5. La síntesis neoclásica y el monetarismo **152**
Introducción . . . 152
1. La síntesis neoclásica . . . 154
1.1. Principales rasgos . . . 154
1.2. Análisis del lado de la demanda . . . 155
1.3. Análisis del lado de la oferta . . . 163
1.4. Conclusiones . . . 167
2. Monetaristas . . . 168
2.1. Principales rasgos . . . 168
2.2. La crítica por el lado de la demanda . . . 169
2.3. La crítica por el lado de la oferta . . . 178
2.4. Conclusiones . . . 181
3. Otros desarrollos posteriores . . . 183
3.1. La nueva macroeconomía clásica . . . 183
3.2. La nueva economía keynesiana de primera generación . . . 183
3.3. La nueva economía keynesiana de segunda generación . . . 183
Conclusiones . . . 184
Anexo. Síntesis por parámetros . . . 185

A6. La nueva macroeconomía clásica **186**
Introducción . . . 186
1. Fundamentos del pensamiento de la NMC . . . 188
1.1. Principales rasgos . . . 188
1.2. La hipótesis de las expectativas racionales (Muth, 1961) . . . 188
1.3. La crítica de Lucas (1976) . . . 189
1.4. Conclusiones . . . 191
2. Aplicaciones y desarrollos: estudio del ciclo walrasiano . . . 192
2.1. Modelo OA-DA con información imperfecta (Lucas, 1972; 1973; 1975) . . . 192
2.2. Ciclos debidos a *shocks* reales no anticipados (Kydland y Prescott, 1982; Long y Plosser, 1983) . . . 196
2.3. Comparación y legado: los modelos DSGE . . . 200
2.4. Conclusiones . . . 201
3. Implicaciones del pensamiento de la NMC . . . 202
3.1. Implicaciones de política monetaria . . . 202
3.2. Implicaciones de política fiscal . . . 208
3.3. Críticas al pensamiento de la NMC . . . 211
3.4. La nueva síntesis neoclásica . . . 212
Conclusiones . . . 216
Anexos . . . 217
A.1. La parábola de las islas de Phelps y la curva de oferta de Lucas . . . 217
A.2. La curva de Phillips con HER . . . 217

A7. La nueva economía keynesiana **219**
Introducción . . . 219

1. Fundamentos del pensamiento neokeynesiano 221
1.1. Principales rasgos 221
1.2. Teorías sobre rigideces reales 222
1.3. Teorías sobre rigideces nominales 229
1.4. Análisis dinámico de las rigideces 233
1.5. Conclusiones 237
2. Aplicaciones y desarrollos: el estudio del ciclo con rigideces 238
2.1. Modelo de la NEK en economía cerrada: IS-PC-MR 238
2.2. Extensiones al modelo 244
2.3. Conclusiones 246
3. Implicaciones del pensamiento de la NEK 247
3.1. Implicaciones de política monetaria 247
3.2. Implicaciones de política fiscal 249
3.3. Implicaciones de política de empleo: políticas de demanda y estructurales (Layard y Nickell, 1985) 252
3.4. Críticas al pensamiento de la NEK 252
3.5. La nueva síntesis neoclásica 254
Conclusiones 255
Anexo. Microfundamentación del modelo de la NEK 2G 256
A.1. Agentes y equilibrio 256
A.2. Obtención de la IS dinámica 258
A.3. Obtención de la CP-NEK 258

Nota

Este PDF es un extracto de *Lecciones de economía general e internacional* (Lecciones de economía, vol. 1), de Arturo Lorenzo, con el bloque «Pensamiento Económico» (temas A1-A7). Se distribuye gratis bajo licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 (CC BY-NC-SA 4.0).

Versión de noviembre de 2026, generada a partir de la edición web. La edición completa y actualizada está en https://librovol1.arturolorenzo.com.

A1. Objeto y método de la ciencia económica

Objeto y métodos de la ciencia económica. Cuestiones y debates actuales, con especial referencia a la economía conductual.

Introducción

John Stuart Mill abría su ensayo sobre la definición de la economía política (1836) con una observación que sigue vigente: "la definición de una ciencia, casi sin excepción, no ha precedido a la creación de la ciencia misma, sino que la ha seguido". La comparaba con la muralla de una ciudad, que se levanta para circundar un conjunto de edificios ya existente y no para acoger los que puedan surgir después.

La economía no es una excepción. La reflexión sobre la producción, el cambio y el dinero es muy anterior a la disciplina[1][2], pero solo cuando la síntesis de Smith y de los clásicos ofrece un cuerpo de doctrina que delimitar, a finales del siglo XVIII, empieza el debate sobre su objeto y su método, que sigue abierto[3].

Ese debate no es solo académico, porque el objeto y el método que aquí se fijan recorren todo el libro. La microeconomía (A08-A24) y la economía laboral (A25-A28) aplican a consumidores, empresas y trabajadores el análisis de la elección en condiciones de escasez, la macroeconomía (A29-A45) lo traslada a los agregados, y la Parte B lo extiende a la empresa, las finanzas, el comercio y las instituciones internacionales. En todos los temas se repite el mismo procedimiento, que consiste en construir un **modelo** a partir de unos supuestos, deducir sus implicaciones y contrastarlas con los datos. Las críticas que este tema recoge al final, de la economía conductual a la macroeconomía posterior a 2008, reaparecen allí donde esos supuestos fallan.

Hoy existe un amplio consenso en considerar la economía una ciencia, aunque queda por precisar de qué tipo de ciencia se trata.

[1] Más allá de los mercantilistas y de los fisiócratas, predecesores inmediatos de los clásicos, pueden citarse Hesíodo, *Los trabajos y los días* (c. 700 a. C.), el *Guanzi* (atribuido a Guan Zhong, s. VII a. C.), Platón, *La República* (c. 375 a. C.), Jenofonte, *Formas y medios de aumentar los ingresos de Atenas* (c. 355 a. C.), Aristóteles, *La política* (c. 330 a. C.), Abu Hamid Al-Ghazali, *Ihya Ulum al Din* (c. 1100) o Santo Tomás de Aquino, *La suma teológica* (1273).

[2] Entre 1500 y 1750 creció notablemente la literatura económica en Europa occidental, obra sobre todo de comerciantes que escribían tratados o panfletos sobre problemas concretos de política económica. Entre 1650 y 1750 empezó a formarse un cuerpo de saber económico.

[3] Schumpeter vincula la consolidación de la economía como disciplina a la "situación clásica" de finales del s. XVIII, en la que *La riqueza de las naciones* (1776) actúa como obra de síntesis (Schumpeter, 1954).

1. *Ciencia social*: con carácter general, se ha considerado que la economía forma parte de las ciencias sociales, en la medida en que examina un problema fundamental de la sociedad, la escasez relativa (apartado 1.2.3)[4].

2. *Relación con las ciencias naturales*: sin embargo, la tesis tradicional de que la economía, a diferencia de las ciencias naturales, no puede realizar experimentos controlados ha sido en ocasiones puesta en entredicho por la economía experimental (apartado 3.1)[5].

Para caracterizar la economía como ciencia hay que determinar antes, como en cualquier otra disciplina, su objeto (apartado 1.1) y su método (apartado 2.1.1). El tema aborda ambas cuestiones en este orden:

- *Objeto*: se busca una definición del objeto de la economía precisa y, a la vez, válida con independencia de la escuela de pensamiento.

- *Método*: se estudian el método científico en general y su aplicación a la economía.

- *Cuestiones y debates actuales*: se repasan las líneas de investigación que discuten supuestos tradicionales de la disciplina, en especial la economía conductual, y el debate sobre el método de la macroeconomía, de la crítica de Lucas a la crisis de 2008.

[4] Históricamente se han utilizado cuatro mecanismos de asignación, no excluyentes entre sí: la fuerza, la tradición, la autoridad (Estado, Iglesia) y el mercado, hoy el principal. Todos ellos determinan quién obtiene los recursos y quién queda sin ellos, por lo que el problema de la escasez plantea también una cuestión de equidad.

[5] Con precedentes desde los años treinta, la economía experimental llega a 1960 con bases sólidas y crece de forma sostenida en los años sesenta (Roth, 1993). Desde los años setenta se consolidan dos líneas, la de V. Smith y C. Plott, centrada en el funcionamiento de los mercados, y la de los psicólogos D. Kahneman y A. Tversky, centrada en la racionalidad de los individuos en el proceso de decisión. Entre sus inconvenientes figura el efecto Hawthorne (los sujetos reaccionan al saberse observados).

1. Objeto de la ciencia económica

Antes de estudiar cómo conoce la economía, conviene precisar qué estudia.

1.1. Concepto de objeto de una disciplina

Toda disciplina científica se identifica, en primer lugar, por aquello que estudia, de modo que la noción de objeto es común a todas las ciencias y previa al estudio de cualquiera de ellas.

Definición: en términos generales, el objeto de una ciencia son los fenómenos que esta estudia. Siguiendo la tradición escolástica, es habitual distinguir entre el *objeto material*, la parcela de la realidad que se estudia, y el *objeto formal*, el aspecto o la perspectiva desde la que se estudia.

Características: para delimitar el objeto de una disciplina conviene tener en cuenta tres rasgos:

- *Acotación*: el objeto es la materia de estudio de una disciplina y no debe confundirse con su objetivo, es decir, con su finalidad.

- *Coincidencia con otras disciplinas*: el objeto material de una ciencia suele solaparse en parte con el de otras, por lo que su delimitación descansa sobre todo en el objeto formal.

- *Carácter histórico y convencional*: la delimitación del objeto se precisa a medida que la disciplina se desarrolla y puede variar de una escuela a otra. Cabe entenderla como una definición nominal y no real, que más que descubrir una esencia fija una convención, y que se juzga por su utilidad para ordenar la investigación.

1.2. Objeto de la economía

Aplicadas a la economía, estas nociones permiten separar lo que comparte con otras ciencias sociales, su objeto material, de lo que le es propio, su objeto formal, que ha cambiado con las escuelas de pensamiento hasta llegar a la definición hoy dominante.

1.2.1. Concepto

Definición: el objeto material de la economía es la actividad humana relacionada con la producción, la distribución, el intercambio y el consumo de bienes y servicios.

Características:

- *Objeto material compartido*: la economía coincide en parte con otras ciencias sociales, como la sociología, la ciencia política, la psicología o la historia, que también estudian la conducta humana y las instituciones.

- *Objeto formal propio*: lo que distingue a la economía no es tanto el fenómeno estudiado como la perspectiva desde la que lo aborda, que hoy es la elección en condiciones de escasez.

1.2.2. Evolución histórica

Históricamente el objeto de la economía ha evolucionado de acuerdo con las inquietudes de diferentes escuelas de pensamiento:

Escuela clásica ortodoxa: formada por autores como Adam Smith o David Ricardo, estuvo especialmente preocupada por el crecimiento, la formación de precios relativos y las fuerzas que alteran la distribución de la renta. Su objeto de estudio giraba en torno a la renta y la riqueza de los países.

La expresión *economía política* aparece en Montchrestien (*Traicté de l'oeconomie politique*, 1615) y la difunde Steuart (1767), para quien la economía política es en el Estado lo que la economía doméstica en la familia. Smith (1776) concibe la economía política como una rama de la ciencia del estadista o del legislador con dos fines, proporcionar al pueblo un ingreso abundante y dotar al Estado de recursos suficientes para los servicios públicos. Sus sucesores, en cambio, la convierten en una ciencia de leyes. Say (1803) la define como el estudio de cómo se forman, distribuyen y consumen las riquezas, y Senior (1836), como la ciencia que trata de la naturaleza, la producción y la distribución de la riqueza. Mill (1836) la restringe a las leyes de los fenómenos sociales que resultan de la acción combinada de los hombres para producir riqueza, en la medida en que no los modifica la búsqueda de otro fin. Al aislar un motivo, Mill ya define la disciplina por un aspecto de la conducta y no solo por una clase de objetos, como hará Robbins. La definición pasa así de un arte de gobierno a una ciencia positiva.

Escuela clásica heterodoxa[6]: la integraba un grupo heterogéneo de críticos de la ortodoxia clásica, desde dentro (Thomas Robert Malthus, Jean-Charles-Léonard de Sismondi) o desde fuera (Friedrich List, Karl Marx). Muchos de estos autores se centraron en las relaciones sociales, históricas, demográficas y de poder. Para Marx, el objeto es descubrir "la ley económica que rige el movimiento de la sociedad moderna" (*El capital*, prólogo de 1867), y la economía política estudia relaciones sociales de producción propias de un periodo histórico, y no leyes naturales y eternas, como pretendían los clásicos.

Escuela neoclásica: integrada por autores como Léon Walras o Alfred Marshall, mantuvo la visión clásica del mercado competitivo como mecanismo autorregulador. Incorporó, sin embargo, la utilidad marginal a la teoría del valor y analizó con más profundidad los mercados y los sistemas de precios[7]. El objeto se desplaza así de las clases y las naciones a los agentes individuales que actúan en los mercados y a las condiciones óptimas de asignación de recursos.

Marshall (1890), que populariza el término *economics* frente a *political economy*, define la disciplina como el estudio de la humanidad en los asuntos ordinarios de la vida. En esa definición *material*, que comparte Cannan y que Robbins criticará, la economía abarca la parte de la acción individual y social más relacionada con la obtención y el uso de los requisitos materiales del bienestar.

La otra rama marginalista prepara la definición actual. Menger (1871) llama actividad económica a la *economización* de los bienes cuya cantidad disponible es inferior a las necesidades, y Wicksteed (1910) y Mises (1922) extienden la lógica de la elección a toda administración de recursos. Robbins parte de ellos.

[6] Según Landreth y Colander, el pensamiento heterodoxo abarca toda corriente que desafía a la corriente principal (p. ej. austriacos, institucionalistas, poskeynesianos o radicales).

[7] En especial, mediante la estática comparativa, que permite entender de qué dependen los precios relativos, qué clases y volúmenes de bienes de consumo se producen, qué clases y dimensiones de empresas se emplean y cómo se establece la distribución personal y funcional de la renta.

Escuelas macroeconómicas: la keynesiana, la nueva macroeconomía clásica, la nueva economía keynesiana y la nueva síntesis neoclásica (Goodfriend y King, 1997) no redefinen la disciplina, sino su problema central. Así, Keynes desplaza la atención de la asignación de recursos plenamente empleados al nivel de producción y empleo, y las escuelas posteriores discuten si las fluctuaciones son desequilibrios o respuestas óptimas a perturbaciones (temas A04-A07). El propio Keynes concibe la economía como "una ciencia moral", que consiste en "pensar en términos de modelos", unida al "arte de elegir los modelos relevantes para el mundo contemporáneo" (cartas a Harrod, julio de 1938).

1.2.3. Concepto actual

La cuestión es si existe un objeto de la economía universal y atemporal, común a las distintas fases de la historia de la teoría económica e independiente de los autores.

Definición: Lionel Robbins propuso en 1932[8] una definición que ha gozado de amplia aceptación entre los economistas: "la economía es aquella ciencia que estudia la conducta humana como una relación entre fines y medios escasos que tienen usos alternativos".

Robbins presenta su definición como *analítica* y no *clasificatoria*, porque no selecciona un tipo de conducta, como la orientada al bienestar material de Marshall, sino un aspecto de toda conducta, la forma que le impone la escasez. Dicho de otro modo, define la economía por su objeto formal y no por su objeto material. De ahí su neutralidad, pues "la economía es enteramente neutral entre fines", lo que la separa de la ética. La definición tardó en imponerse, y los manuales no la adoptan de forma general hasta los años sesenta (Backhouse y Medema, 2009). Los manuales posteriores la han hecho suya y la han extendido de la elección individual a la sociedad. Así, Samuelson y Nordhaus (2010) definen la economía como el estudio de cómo las sociedades utilizan recursos escasos para producir bienes valiosos y distribuirlos entre las personas. Para Robbins, la conducta adquiere un aspecto económico cuando hay fines múltiples y de distinta importancia, y tiempo y medios limitados y susceptibles de usos alternativos. No hace falta suponer deseos ilimitados, como suelen decir los manuales, porque basta con que los medios no alcancen para todos los fines a la vez.

Existe, por tanto, cierto consenso en considerar la *escasez relativa* el principal tema del pensamiento económico moderno, y en ella se origina el estudio de los sistemas económicos e institucionales como mecanismos de asignación y distribución de los recursos.

Características: la concepción moderna del objeto de la economía presenta dos rasgos:

- *Generalidad*: se abstrae del contexto histórico de cada autor y ha resultado coherente con la expansión posterior de las fronteras del conocimiento económico.

- *Universalidad*: no depende de un tipo de bienes ni de un sistema económico, de modo que vale para cualquier elección con medios escasos, en una economía de mercado, planificada o doméstica.

La definición de Robbins ha recibido cuatro objeciones:

- *Amplitud excesiva*: si la economía es el estudio de toda elección entre fines y medios escasos, nada la distingue de otras ciencias por su objeto material, sino solo por su objeto formal, la perspectiva desde la que aborda cualquier conducta. Esa amplitud ha permitido el llamado *imperialismo económico*, la extensión del análisis económico a la familia, el delito o la política (Becker, 1976).

[8] Lionel Robbins, *An Essay on the Nature and Significance of Economic Science* (1932).

- *Recursos ociosos*: la definición pone el acento en la asignación de recursos que se suponen plenamente empleados. Sin embargo, con desempleo involuntario, como el keynesiano, el problema no es asignar medios escasos, sino movilizar los ociosos, y la escasez deja de ser la restricción relevante.

- *Intercambio e instituciones*: Buchanan (*What Should Economists Do?*, 1964) propone centrar la economía en el intercambio y en las instituciones del mercado (la *catalláctica*) y no en la asignación, que sería un problema técnico de un decisor único.

- *Fines y medios dados*: la definición toma como datos los fines y los medios, y deja fuera la formación de las preferencias y el descubrimiento de medios nuevos, que para Kirzner (1973) es la función del empresario y para la economía evolutiva (Nelson y Winter, 1982), el motor del cambio económico.

De ahí han surgido definiciones alternativas, como la economía como estudio del *aprovisionamiento* de la vida humana, propia del institucionalismo y de la economía feminista (Nelson, 1993), o la respuesta pragmática atribuida a Jacob Viner, para quien la economía "es lo que hacen los economistas". Esta última refleja la distancia entre las definiciones formales y la práctica de los economistas (Backhouse y Medema, 2009).

Microeconomía y macroeconomía: en el pensamiento económico moderno, el problema de la escasez relativa se divide en "subobjetos" que, tradicionalmente, se reparten entre estos dos cuerpos teóricos[9]:

1. *Microeconomía*: estudia el comportamiento de agentes económicos individuales (p. ej. consumidores, empresas o sector público) y su interacción en mercados o industrias concretos (análisis de equilibrio parcial) o en varios mercados o industrias interdependientes (análisis de equilibrio general).

2. *Macroeconomía*: estudia el funcionamiento de la economía en su conjunto y los fenómenos agregados (p. ej. inflación o crecimiento) de países o bloques regionales.

No obstante, la frontera entre ambos cuerpos teóricos es cada vez más difusa, sobre todo porque la macroeconomía descansa en una microfundamentación cada vez más sofisticada.

Economía positiva y economía normativa: por último, el carácter de ciencia social de la economía la liga a cuestiones éticas de difícil respuesta. Por ello, su estudio admite dos enfoques, cuya distinción condiciona también el método (apartado 2.2)[10][11]:

1. *Economía positiva*: describe, explica y predice los hechos económicos. J. N. Keynes (1891) la define como un cuerpo de conocimiento sistematizado sobre lo que es. Weber (1904, 1917) defiende la *neutralidad valorativa* de las ciencias sociales, y Robbins (1932) sostiene que la economía y la ética "no están en el mismo plano del discurso". Friedman (1953) la considera, en principio, independiente de cualquier juicio normativo y juzga sus teorías por la precisión, el alcance y la conformidad con la experiencia de sus predicciones. Su fundamento lógico es la *ley de Hume* (Hume, 1739-1740), según la cual de enunciados sobre lo que es no puede deducirse un enunciado sobre lo que debe ser.

[9] La distinción suele atribuirse a Ragnar Frisch, que en 1933 distinguió entre análisis micro y macrodinámico.

[10] Senior (1836) y Mill (1836) ya distinguían entre la ciencia de la economía política, que describe y explica, y su "arte", que prescribe.

[11] Senior y Mill usan esta distinción para delimitar la ciencia, que para ellos comprende solo la economía positiva, y el economista, en palabras de Senior, no puede añadir a sus conclusiones "una sola sílaba de consejo".

2. *Economía normativa*: estudia lo que debe ser, es decir, los criterios para valorar situaciones y políticas, que descansan en juicios de valor (eficiencia, equidad, libertad). No es independiente de la positiva, porque toda recomendación combina un objetivo con una predicción sobre los efectos de cada medio. J. N. Keynes añadía un tercer ámbito, el *arte* de la economía, el sistema de reglas para alcanzar un fin dado, antecedente de la actual política económica.

La frontera entre ambos enfoques es menos nítida de lo que sugiere la ley de Hume, y la literatura la ha precisado con tres distinciones:

- *Juicios metodológicos*: Nagel (1961) distingue los juicios de valor que *caracterizan*, es decir, que estiman en qué grado algo cumple un criterio, de los que *aprecian*, es decir, que lo declaran deseable. Toda ciencia necesita juicios metodológicos, sobre qué problemas estudiar o qué evidencia basta, pero eso no la vuelve normativa (Blaug, 1992).

- *Juicios básicos y no básicos*: según Sen (1970a), un juicio de valor es no básico si su validez depende de circunstancias de hecho, de modo que la evidencia empírica puede obligar a revisarlo. Buena parte del desacuerdo sobre políticas es de este tipo.

- *Valores en los conceptos*: Myrdal (1930, 1953) sostuvo que los juicios de valor intervienen ya en la elección de los problemas y los conceptos, y Putnam (2002), que conceptos como el bienestar o la pobreza son a la vez descriptivos y valorativos. Para Putnam se desmorona la dicotomía entre hechos y valores, no la distinción.

Por ello, la exigencia habitual, que ya formuló Myrdal, no es eliminar los juicios de valor, sino hacerlos explícitos.

1.3. Conclusiones

El objeto de la economía no ha sido fijo. Cada escuela lo ha delimitado según sus preocupaciones, desde la riqueza de las naciones y la distribución entre clases de los clásicos hasta la asignación de recursos entre agentes individuales de los neoclásicos y los agregados de la macroeconomía keynesiana. Aun así, la definición de Robbins goza hoy de amplia aceptación porque es lo bastante general para abarcar esos enfoques y las extensiones posteriores.

Esa misma generalidad es también su principal debilidad. Al definir la economía por una perspectiva y no por una parcela de la realidad, la definición de Robbins deja abiertas las fronteras de la disciplina con otras ciencias sociales y encaja mal con la existencia de recursos ociosos. Además, relega las instituciones del intercambio y toma como dados los fines y los medios. De ahí que convivan con ella definiciones alternativas. Las divisiones del objeto ordenan su estudio, aunque la frontera entre microeconomía y macroeconomía es cada vez más difusa, y la distinción entre economía positiva y normativa, que muchos entienden hoy como una distinción y no como una dicotomía, enlaza con el método (apartado 2).

2. Método de la ciencia económica

Definido el objeto, este apartado estudia primero el método científico en general, la evolución de la reflexión sobre él y su concepción actual, y después, en el mismo orden, el método de la economía.

2.1. Método científico

Antes de pasar a la economía, este apartado define el método científico y lo distingue de las técnicas y de la metodología, y repasa después las corrientes de la filosofía de la ciencia que lo han debatido, del racionalismo de Descartes al anarquismo de Feyerabend, para cerrar con su concepción actual.

2.1.1. Concepto

Definición: el *método científico* (del griego *méthodos*, "camino para llegar a un fin") es el procedimiento ordenado con el que se obtiene y se justifica el conocimiento científico. Mientras que el objeto se refiere a qué aspecto de la realidad se analiza (apartado 1.1), el método se refiere a cómo se analiza. Por ello, busca reglas de carácter general[12] que permitan pasar de un problema a enunciados contrastados con la experiencia.

Conviene distinguir tres niveles. Por un lado, están las *técnicas*, que son los instrumentos concretos de cada disciplina, como la encuesta, la regresión o el experimento. Por otro lado, está el *método*, que es la lógica común que ordena el uso de esas herramientas. Finalmente, está la *metodología*, que es el estudio crítico de los métodos, es decir, la filosofía de la ciencia aplicada a una disciplina, que examina con qué criterios se aceptan o se rechazan sus teorías (Blaug, 1992). Dentro de ella, Reichenbach (1938) separa el *contexto de descubrimiento*, cómo se llega a concebir una hipótesis, del *contexto de justificación*, cómo se decide si es aceptable. Las metodologías que se exponen a continuación se ocupan sobre todo del segundo.

El razonamiento científico combina tres formas de inferencia:

- *Deducción*: obtiene conclusiones que se siguen necesariamente de las premisas, de modo que preserva la verdad, pero no amplía lo que estas ya contienen. El racionalismo privilegia este razonamiento.

- *Inducción*: generaliza a partir de casos observados, de modo que amplía el conocimiento, pero sin certeza. El empirismo privilegia este razonamiento.

- *Abducción*: infiere la hipótesis que, de ser cierta, explicaría un hecho sorprendente. Charles S. Peirce la llamó primero *hipótesis* (1878) y después *abducción* o *retroducción*, y la situó en la fase en que se generan las conjeturas, que después la deducción desarrolla y la inducción somete a contraste.

La reconstrucción habitual de cómo se combinan estas inferencias es el *método hipotético-deductivo*, que Hempel (1966) presenta como el procedimiento propio de las ciencias empíricas. Este consta de cuatro pasos:

1. *Problema*: la observación plantea un hecho que las teorías disponibles no explican.

[12] Las reglas son válidas independientemente del sujeto que las aplique.

2. *Hipótesis*: se propone una conjetura explicativa, por abducción o por cualquier otra vía, ya que no existe una regla para inventarla.

3. *Deducción de consecuencias*: de la hipótesis, junto con hipótesis auxiliares y condiciones iniciales, se deducen predicciones observables.

4. *Contrastación*: las predicciones se comparan con la observación o el experimento. Si fallan, la hipótesis, o alguna de las auxiliares, queda refutada. En cambio, si se cumplen, la hipótesis solo queda confirmada de forma provisional y se somete a nuevas pruebas.

Según el modelo nomológico-deductivo de Hempel y Oppenheim (1948), el mismo esquema sirve para explicar, porque explicar un hecho consiste en deducirlo de leyes generales y condiciones iniciales, de modo que explicación y predicción comparten la misma estructura lógica (*tesis de la simetría*). Esta idea es afín al peso que Friedman concede a la predicción en la evaluación de las teorías económicas.

De este esquema se derivan las nociones con las que la filosofía de la ciencia ha debatido el método:

- *Contrastación*: confrontación de las consecuencias de una hipótesis con la evidencia.

- *Verificación*: prueba concluyente de la verdad de un enunciado mediante la experiencia. Es posible para enunciados singulares, pero no para leyes universales, que se refieren a un número ilimitado de casos.

- *Confirmación*: aumento del apoyo empírico de una hipótesis cuando se cumplen sus predicciones. Es gradual y nunca definitiva.

- *Falsación*: refutación de una hipótesis porque una de sus predicciones resulta falsa. Por la forma lógica del *modus tollens*, una sola observación contraria basta para refutar un enunciado universal.

- *Demarcación*: criterio que separa la ciencia de lo que no lo es, como la verificabilidad del positivismo lógico o la falsabilidad de Popper.

2.1.2. Evolución histórica

El método científico ha estado sujeto a evolución. En este sentido, Aristóteles suele considerarse autor del primer tratamiento sistemático del conocimiento científico[13]. El filósofo griego concibe la ciencia como demostración, una deducción silogística a partir de primeros principios, a los que se llega por inducción (*epagogé*) desde la experiencia. En su obra están ya, por tanto, las dos formas de razonamiento que vertebran el debate posterior.

Sin embargo, el método científico moderno nace con la *revolución científica* del s. XVII, que une la experimentación y la matemática en la obra de Galileo y culmina en los *Principia* de Newton (1687). Su reflexión filosófica arranca con Bacon (*Novum organum*, 1620) y Descartes (*Discurso del método*, 1637)[14]. A partir de ahí pueden distinguirse siete grandes corrientes, en orden aproximadamente cronológico.

[13] Este estudio se recoge, principalmente, en los tratados lógicos reunidos en el *Órganon* ("instrumento"), en particular los *Primeros analíticos* y los *Segundos analíticos*.

[14] En el *Discurso* Descartes formula el "pienso, luego existo" y parte de la duda metódica (dudar de todo) para liberar la razón de nociones preconcebidas.

Racionalismo y deductivismo (René Descartes, 1596-1650): el método de Descartes es conocido como método deductivo o axiomático-deductivo.

- *Definición*: el método deductivo aplica reglas lógicas a un conjunto de axiomas de partida (verdades que se tienen por evidentes en sí mismas) para obtener leyes universales que se enuncian en forma de teoremas.

- *Proceso*: bajo este método, la ciencia avanza por medio del *ejercicio de la razón*.

- *Valoración*: como toda deducción, el método deductivo no amplía el contenido empírico de sus premisas. Por eso su valor depende de la verdad de los axiomas, y en las ciencias sociales es difícil plantear axiomas que no generen disputa.

Empirismo e inductivismo (Francis Bacon, 1561-1626; David Hume, 1711-1776): frente al método deductivo surge también en el s. XVII el método inductivo del empirismo británico[15].

- *Definición*: el método inductivo, inverso al deductivo[16], parte de la observación de hechos, descubre regularidades y las generaliza en leyes universales y teorías.

- *Proceso*: para el inductivismo, la ciencia parte de la *observación sin prejuicios* y avanza generalizando a partir de ella, aunque Bacon rechazó la simple enumeración de casos y propuso una inducción eliminativa[17].

- *Valoración*: ningún número finito de observaciones justifica lógicamente un enunciado universal. Que una regularidad se haya cumplido hasta ahora no garantiza que se cumpla en la siguiente observación, y apelar a que la inducción ha funcionado en el pasado es a su vez un razonamiento inductivo, dificultad conocida como *problema de la inducción* (Hume, 1739).

Criticismo kantiano (Immanuel Kant, 1724-1804): Kant intenta reconciliar ambas posturas en la *Crítica de la razón pura* (1781).

- *Definición*: el criticismo examina la propia facultad de conocer para fijar sus condiciones y sus límites. Su tesis es que todo conocimiento comienza con la experiencia, pero no todo procede de ella.

- *Proceso*: la mente aporta formas *a priori* (el espacio, el tiempo y las categorías del entendimiento, como la causalidad) que ordenan los datos *a posteriori* de los sentidos, de modo que "los pensamientos sin contenido son vacíos y las intuiciones sin conceptos son ciegas". Ello permite los *juicios sintéticos a priori*, que amplían el conocimiento y son, a la vez, universales y necesarios. Estos juicios están detrás del avance científico.

- *Valoración*: el positivismo lógico rechazó los juicios sintéticos *a priori*, porque la geometría no euclídea y la relatividad general habían mostrado que la geometría euclídea, que Kant tenía por sintética *a priori*, no es una verdad necesaria sobre el espacio físico. Sin embargo, la idea kantiana de que los conceptos previos ordenan la experiencia anticipa la *carga teórica de la observación* que subrayarán después Hanson y Kuhn.

[15] El empirismo británico incluye también a John Locke y George Berkeley.

[16] En el método deductivo, se comenzará planteando enunciados de carácter general a partir de los cuales se inferirán enunciados de carácter más concreto o particular (Todos los hombres son mortales. Keynes es un hombre. Luego Keynes es mortal.), mientras que en el método inductivo se irá de lo particular a lo general (El cisne 1 es blanco, el cisne 2 es blanco… el cisne 100 es blanco; luego todos los cisnes son blancos).

[17] El método baconiano, más elaborado que esta síntesis, se expone en el *Novum organum* (1620), concebido como alternativa al *Órganon* aristotélico. Mill (*Sistema de lógica*, 1843) sistematizó después cinco métodos de inferencia inductiva (concordancia, diferencia, conjunto de concordancia y diferencia, residuos y variaciones concomitantes), antecedente de la inferencia causal moderna.

Positivismo y positivismo lógico (Auguste Comte, 1798-1857; Círculo de Viena, 1924-1936): el término procede de Auguste Comte (*Curso de filosofía positiva*, 1830-1842).

▪ *Definición*: para Comte, el conocimiento válido es el que se apoya en la observación de los hechos y en las leyes que los relacionan, frente a las explicaciones teológicas o metafísicas. En el s. XX, el positivismo lógico o empirismo lógico del Círculo de Viena (Schlick, Carnap, Neurath, entre otros) une ese empirismo a la lógica formal de Frege y Russell. Su tesis central, el *criterio verificacionista del significado*, sostiene que un enunciado que no es analítico (verdadero por su sola forma lógica o por el significado de sus términos) solo tiene significado si puede verificarse empíricamente. Por eso, los enunciados de la metafísica, inverificables, plantean pseudoproblemas.

▪ *Proceso*: la ciencia avanza por *verificación* y, en la versión posterior del propio Carnap (1936-1937), por *confirmación gradual*. Las teorías se reconstruyen como sistemas axiomáticos cuyos términos teóricos se conectan con la observación mediante reglas de correspondencia, y explicar un hecho es deducirlo de leyes generales (Hempel y Oppenheim, 1948). Esta *concepción heredada* (*received view*) dominó la filosofía de la ciencia hasta los años sesenta.

▪ *Valoración*: el criterio tropieza con tres dificultades. Por un lado, las leyes universales no pueden verificarse de forma concluyente, por el problema de la inducción, de modo que el criterio privaría de significado a las propias leyes científicas. Además, el criterio no es analítico ni verificable, por lo que no supera su propia prueba. Por último, Quine (*Two Dogmas of Empiricism*, 1951) negó que exista una separación nítida entre enunciados analíticos y sintéticos y sostuvo que los enunciados se contrastan en bloque y no uno a uno. La primera dificultad es el punto de partida de Popper.

Racionalismo crítico o falsacionismo (Karl Popper, 1902-1994): el problema de la inducción es uno de los principales argumentos con los que Popper defiende en *La lógica de la investigación científica* (1934) un nuevo método, el deductivo contrastable[18].

▪ *Definición*: el método parte de hipótesis formuladas como proposiciones universales no restringidas y contrastables empíricamente. Su fundamento es la asimetría entre verificación y falsación.

Popper propone la falsabilidad para distinguir lo que es ciencia de lo que no lo es. Es decir, establece un *criterio de demarcación* de la ciencia. Fuera de la ciencia sitúa la metafísica y el psicoanálisis, por no ser empíricamente falsables, y el marxismo, cuyos seguidores lo volvieron irrefutable reinterpretando sus predicciones fallidas[19]. A diferencia del verificacionismo, la falsabilidad no es un criterio de significado, porque un enunciado metafísico puede tener sentido e incluso inspirar teorías científicas.

▪ *Proceso*: en este contexto, la ciencia avanza por *conjeturas y refutaciones* (Popper, 1963). Se proponen hipótesis audaces, de gran contenido empírico, y se someten a las pruebas más severas posibles. La falsación permite desechar las que fallan, mientras que las que resisten quedan *corroboradas*, que no es lo mismo que verificadas, porque siguen expuestas a refutaciones futuras.

▪ *Valoración*: la falsación tropieza con tres problemas:

[18] Popper es inusual para un filósofo contemporáneo en la medida en que acepta la crítica de Hume de la inducción y va más allá al defender que la inducción no ha sido nunca utilizada en la ciencia. No obstante, no se alinea con el escepticismo asociado a Hume (Popper, 1963, cap. 1).

[19] Como ejemplo de teoría científica, Popper cita la relatividad general de Einstein, que formulaba predicciones arriesgadas y contrastables, como la desviación de la luz por el Sol, comprobada por Eddington en el eclipse de 1919.

- *Tesis de Duhem-Quine*: no es posible falsar de manera concluyente una hipótesis científica concreta, porque se contrasta siempre junto con un conjunto de hipótesis auxiliares implícitas. Por ello, nunca se sabe con certeza si la refutación obedece a fallos de la hipótesis principal o de las auxiliares (Duhem, 1906; Quine, 1951). Popper responde con una regla metodológica que prohíbe introducir supuestos auxiliares *ad hoc* para evitar la refutación, las *estratagemas convencionalistas* (Popper, 1934), que Albert llamó después *inmunizadoras*.

- *Falsacionismo ingenuo*: Lakatos (1970) distingue tres versiones del falsacionismo. El *dogmático* tiene por concluyente toda refutación empírica. El *metodológico ingenuo* acepta por convención los enunciados de observación, pero descarta toda teoría refutada. El *sofisticado* solo abandona una teoría cuando otra explica lo mismo y predice, además, hechos nuevos corroborados. La lectura ingenua, que Lakatos atribuye a una parte de la obra de Popper, haría demasiado errática la evolución de la ciencia.

- *Hipótesis probabilísticas*: habituales en economía, no son falsables en sentido estricto, porque ninguna muestra finita es lógicamente incompatible con ellas. Popper (1934) propone tratarlas como refutadas cuando los datos son muy improbables bajo la hipótesis, una decisión metodológica análoga a la de los contrastes estadísticos de Fisher y de Neyman y Pearson, desarrollados de forma independiente.

Nueva filosofía de la ciencia o giro historicista (Thomas Kuhn, 1922-1996; Imre Lakatos, 1922-1974): tras Popper, la filosofía de la ciencia atiende a cómo procede de hecho la ciencia en su historia, y no solo a cómo debería proceder, corriente llamada también pospositivismo.

▪ *Definición*: la unidad de análisis deja de ser la teoría aislada y pasa a ser una estructura más amplia, que cada autor concibe a su manera:

- *Kuhn*: en *La estructura de las revoluciones científicas* (1962) sostiene que las teorías resisten al cambio más de lo que supone el falsacionismo, porque se integran en *paradigmas*. Los paradigmas son las realizaciones científicas universalmente reconocidas que, durante cierto tiempo, proporcionan modelos de problemas y de soluciones a una comunidad científica. Cada uno define qué problemas son legítimos, qué cuenta como solución y qué métodos son aceptables. Más que proponer un método, Kuhn trata de explicar cómo opera y evoluciona en realidad la ciencia.

- *Lakatos*: en *La metodología de los programas de investigación científica* (1978), que recoge póstumamente su ensayo de 1970, mantiene, a diferencia de Kuhn, un criterio normativo para evaluar teorías. Su unidad de evaluación, en cambio, es el *programa de investigación*, una sucesión de teorías que comparten un núcleo y un cinturón protector[20].
  - *Núcleo duro* (*hard core*): hipótesis básicas del programa, que se declaran irrefutables por decisión metodológica (la *heurística negativa* prohíbe dirigir contra ellas la refutación).
  - *Cinturón protector* (*protective belt*): hipótesis auxiliares refutables, que soportan las contrastaciones empíricas y se reajustan para explicar las discrepancias observadas.

▪ *Proceso*: el cambio científico consiste en la *sustitución de esas estructuras*, discontinua en Kuhn y gradual en Lakatos:

[20] Junto a la heurística negativa, el programa se guía por una *heurística positiva*, que indica cómo desarrollar y modificar el cinturón protector, es decir, qué problemas abordar y qué variantes refutables construir.

- *Kuhn*: la ciencia avanza por sustitución de paradigmas[21], con etapas de elevada estabilidad y otras de profundo cambio. Antes del primer paradigma, la disciplina atraviesa un periodo *preparadigmático*, con escuelas rivales que no comparten problemas ni criterios. Después, el proceso recorre dos etapas que se repiten:
    - *Ciencia normal*: los científicos resuelven enigmas dentro del paradigma vigente. De la contrastación empírica surgen discrepancias que, si no se resuelven, cristalizan en anomalías, hechos que el paradigma no es capaz de explicar. Cuando las anomalías se acumulan y resisten, el paradigma entra en *crisis*.
    - *Ciencia revolucionaria*: durante la crisis compiten el paradigma antiguo y uno nuevo, que ya ha ganado a una parte de la comunidad porque resuelve las anomalías. Si el nuevo se impone, se produce una *revolución científica* que da lugar a una nueva ciencia normal. Para Kuhn, la elección entre paradigmas no se decide solo por la lógica y la evidencia, sino también por la persuasión, como una conversión de la comunidad.

- *Lakatos*: el paso de un programa de investigación a otro es gradual y comparativo. Su criterio de demarcación es que solo son científicos los programas progresivos, los únicos que hacen avanzar la ciencia. Los programas pueden ser:
    - *Teóricamente progresivos*: sus adaptaciones permiten predecir hechos nuevos, hasta entonces inesperados.
    - *Empíricamente progresivos*: además, parte de esas predicciones nuevas se ve corroborada por la evidencia.
    - *Degenerativos*: sus revisiones incorporan de forma creciente supuestos *ad hoc* cuyo único propósito es blindar sus teorías frente a los resultados adversos.

▪ *Valoración*: ambos explican mejor que el falsacionismo por qué la comunidad científica no abandona una teoría ante la primera refutación, pero reciben críticas distintas:

- *Kuhn*: se le reprochan tres limitaciones. Primera, la vaguedad del concepto de paradigma, del que Masterman (1970) contó una veintena de acepciones en la obra de Kuhn. Segunda, la *inconmensurabilidad* entre paradigmas sucesivos, que deja sin un criterio racional para preferir uno a otro y, para Lakatos, convierte el cambio científico en una cuestión de "psicología de masas". Tercera, Kuhn no especifica cómo se pasa de una mera acumulación de anomalías a una ruptura efectiva y sustitución de paradigmas.

- *Lakatos*: su metodología, a medio camino entre Popper y Kuhn, no establece, según Feyerabend, cuándo debe abandonarse un programa degenerativo, que siempre puede recuperarse en el futuro. A falta de ese criterio, la distinción entre programas progresivos y degenerativos no dicta ninguna regla al científico, de modo que, para Feyerabend, se trata de un "anarquismo disfrazado".

Anarquismo epistemológico (Paul Feyerabend, 1924-1994): finalmente, Feyerabend invierte por completo los términos del debate metodológico en su obra *Contra el método* (1975).

▪ *Definición*: Feyerabend sostiene que ningún método de principios firmes e inamovibles resiste el contraste con la historia de la ciencia, porque no hay una sola regla metodológica que no se haya vulnerado en algún momento.

[21] Por la amplitud del concepto de paradigma, Kuhn lo sustituyó después (posdata de 1969) por el de *matriz disciplinaria*, el conjunto de generalizaciones simbólicas, modelos, valores y ejemplares compartidos por los miembros de una comunidad científica. Reservó el término paradigma para los *ejemplares*, las soluciones concretas de problemas que sirven de modelo.

Popper y Kuhn difieren en su actitud ante la contrastación empírica. Popper exige para abandonar una teoría un efecto reproducible que la contradiga, no una mera discrepancia, y concibe la ciencia como una revolución permanente, frente a los saltos discontinuos de Kuhn.

- *Proceso*: para Feyerabend, el único principio que no inhibe el progreso científico es "*todo vale*". No pretende con ello abolir las reglas, sino evitar que la investigación se supedite a unas concretas, porque ninguna es infalible.

- *Valoración*: se le reconoce que evita el dogmatismo metodológico, pero menosprecia patrones de descubrimiento que han sido eficaces en la historia de la ciencia.

2.1.3. Concepto actual

Definición: en su acepción actual, el método científico es un procedimiento sistemático, público y autocorrector, que revisa o abandona las hipótesis que fallan. No existe un único método universalmente aceptado, sino una combinación de principios y prácticas consolidados a partir de las aportaciones anteriores:

- *Formulación de hipótesis falsables*: la influencia de Popper sigue siendo dominante, porque la comunidad científica exige que las hipótesis sean contrastables empíricamente. En la práctica, sin embargo, la falsación estricta ha dado paso a un enfoque próximo al de Lakatos, que evalúa los programas de investigación por su capacidad para generar predicciones nuevas y corroboradas.

- *Revisión por pares* (*peer review*): expertos independientes evalúan de forma anónima los trabajos antes de su publicación. Este filtro, aunque imperfecto, es el principal control de calidad de la producción científica. Además, institucionaliza el *escepticismo organizado*, una de las normas del ethos científico de Merton (1942).

- *Replicabilidad*: investigadores independientes han de poder reproducir los resultados de un estudio. Sin embargo, desde la década de 2010, la constatación de que muchos resultados publicados no se reproducen, la llamada **crisis de replicación**, ha impulsado reformas de transparencia, registro previo de estudios (*pre-registration*) y acceso abierto a datos y código.

- *Pluralismo metodológico*: más allá de estos principios compartidos, conviven en la actualidad diversos enfoques. Entre ellos, las metodologías bayesianas, que asignan grados de credibilidad a las hipótesis y los actualizan a medida que se incorpora nueva evidencia (Howson y Urbach, 1989), y la metodología experimental, que se estudia en el apartado 3.1.

Características: Gibson (1960) defiende para las ciencias sociales, frente a quienes niegan su carácter científico, cinco características del procedimiento científico:

- *Abstracción*: selecciona los rasgos esenciales del fenómeno estudiado y prescinde de los accesorios.

- *Generalidad*: busca enunciados válidos para una clase de fenómenos, no para casos aislados.

- *Neutralidad*: separa los juicios de hecho de los juicios de valor.

- *Objetividad*: las circunstancias del investigador (su posición social o sus intereses) no le impiden basar sus enunciados en la evidencia, que otros pueden comprobar.

- *Evidencia empírica*: sus enunciados se apoyan en la observación y se someten a contraste con ella (o con el experimento, cuando es posible).

2.2. Método económico

Repasada la evolución de la metodología científica, este apartado examina su aplicación a la ciencia económica con algunos de los conceptos ya introducidos.

2.2.1. Concepto

Definición: el *método de la economía* puede entenderse hoy como la interrelación de cuatro elementos (la realidad económica; las hipótesis, teorías y modelos; la base empírica, y los resultados) unidos por cuatro operaciones: la abstracción, la medición, la aplicación y predicción, y la contrastación. La figura A1.1 las resume.

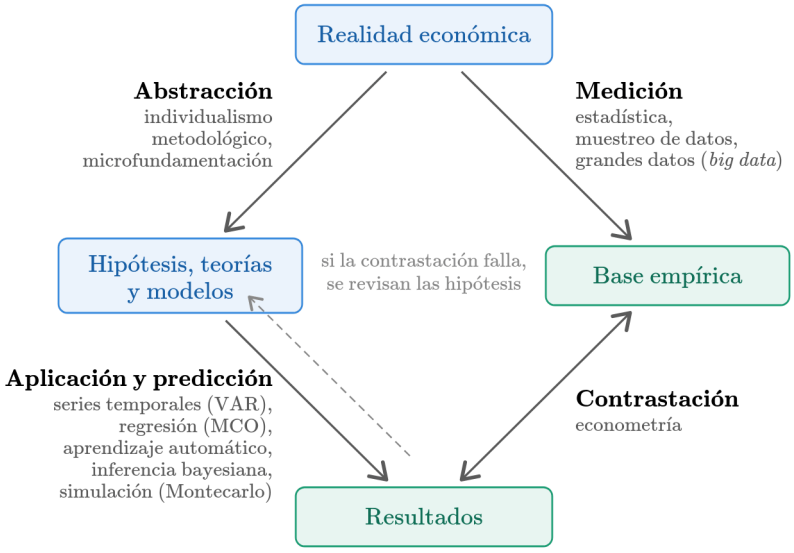

Figura A1.1: Elementos y operaciones del método de la economía. La abstracción lleva de la realidad económica a las hipótesis, teorías y modelos, y la medición, de la realidad a la base empírica. Los modelos se aplican para obtener resultados, que se contrastan con la base empírica, y, si la contrastación falla, se revisan las hipótesis. En azul, el ámbito teórico, y en verde, el empírico.

Las cuatro operaciones forman un ciclo, y si los resultados no superan la contrastación con la base empírica, se revisan las hipótesis.

2.2.2. Evolución histórica

La evolución del método seguido por la economía puede organizarse en cuatro etapas, la primera con una corriente ortodoxa y otra heterodoxa. Entre ellas cambian el modo de razonar, que oscila entre la deducción y la inducción; el lenguaje, que pasa de verbal a gráfico y matemático, y la forma de contrastar las teorías, que va de la verificación de Mill a la econometría y el experimento.

Escuela clásica ortodoxa (s. XVIII-1870): Adam Smith, David Ricardo y John Stuart Mill no siguen un método único y común, pero en sus trabajos predomina el deductivismo a partir

del razonamiento lógico-verbal. Con todo, salvo en Senior y Mill, esta preferencia no es explícita, y los autores también incorporan aspectos inductivos derivados del análisis histórico y contemporáneo:

- *Adam Smith*: combina el razonamiento deductivo sobre la economía de su tiempo con el relato descriptivo de las instituciones contemporáneas e históricas. Une, por tanto, los métodos deductivo e inductivo.

- *David Ricardo*: teórico puro y mayor exponente del deductivismo clásico, enuncia leyes de carácter general sin gran interés por la contrastación empírica[22] y es pionero en la construcción de un modelo económico sistemático, el modelo ricardiano de distribución.

- *John Stuart Mill*: la reflexión metodológica explícita en economía comienza con Nassau Senior (1836) y John Stuart Mill (1836). Mill concibe la economía política como una ciencia abstracta e hipotética, cuyo método es *a priori*. Razona deductivamente a partir de supuestos sobre la conducta (el deseo de riqueza, contrarrestado solo por la aversión al trabajo y el deseo de consumo inmediato), y sus conclusiones solo son ciertas en abstracto, como *tendencias* que otras causas perturbadoras pueden contrarrestar. La verificación *a posteriori* comprueba si en el caso concreto se ha omitido alguna causa relevante, sin contrastar las premisas. Además, Mill describe en su ensayo de 1836 lo que autores posteriores denominaron *homo economicus*, el individuo considerado solo como un ser que desea poseer riqueza, con abstracción de cualquier otra pasión o motivo. Cairnes (1857) y J. N. Keynes (1891) sistematizan después esta posición, para la que las premisas de la teoría son ciertas y la evidencia solo comprueba si la teoría es aplicable al caso. Por ello, Blaug (1992) los llama *verificacionistas*.

Escuela clásica heterodoxa (finales del s. XVIII y s. XIX): junto a estos autores figuran Thomas Robert Malthus y Karl Marx, que siguen, cada uno, una metodología distinta. En sus obras pesa más el análisis histórico y empírico, combinado con el deductivo:

- *Thomas Robert Malthus*: en su teoría de la población, con la que explica las causas de la pobreza, combina el análisis deductivo con el inductivo, derivado del estudio de las fuerzas demográficas de su tiempo.

- *Karl Marx*: su método dialéctico, de raíz hegeliana, y el materialismo histórico, que trata las categorías económicas (como la plusvalía) como propias de un modo de producción históricamente determinado, lo llevan a conclusiones radicalmente distintas de las de los clásicos ortodoxos.

Escuela marginalista, neoclásica e historicista (1870-1910): la Revolución Marginalista de la década de 1870 supone un punto de inflexión metodológico. Se generaliza el uso de herramientas matemáticas y gráficas, y la economía refuerza su carácter cuantitativo y deductivo, no sin fricciones entre marginalistas e historicistas. Durante esta etapa conviven dos enfoques metodológicos:

- *Historicismo*: su método es inductivo, pues se apoya en el análisis histórico y en la recogida de abundante información estadística sobre el desarrollo económico. Surge como respuesta a la popularidad creciente en Europa del *laissez faire* y rechaza explícitamente el ideario de la escuela clásica británica[23]. Su precursor es Friedrich List, y sus exponentes son la escuela histórica

[22] Ello no significa que ignorase los problemas de su época, como los aranceles sobre el cereal importado (Corn Laws) o el conflicto entre terratenientes y capitalistas, pero sostuvo que la teoría era requisito previo de cualquier análisis. Nassau Senior y los seguidores de Ricardo, en cambio, se centraron en exceso en la deducción y dejaron de lado la evidencia empírica.

[23] Con raíces en la filosofía hegeliana y en el romanticismo de Adam Müller, era de corte nacionalista y proteccionista, porque la política económica adecuada para un país industrializado como Inglaterra no lo era para Alemania o EE. UU., todavía eminentemente agrícolas.

alemana antigua (Roscher, Hildebrand, Knies) y, sobre todo, la joven escuela histórica de Gustav Schmoller.

▪ *Escuela marginalista*: recurrió al método deductivo y a modelos abstractos, con técnicas que variaban de un autor a otro:

- *Escuela británica*: W. S. Jevons abogó por un mayor uso de procedimientos estadísticos para establecer relaciones causales entre variables económicas. Por su parte, A. Marshall se apoyó, sobre todo, en el razonamiento gráfico y, subsidiariamente, matemático. Por tanto, los autores británicos combinaron el enfoque inductivo y el deductivo.

- *Escuelas continentales*: se desdoblan en dos:

  - *Escuela austriaca*: C. Menger abogó por recurrir en mayor medida al razonamiento abstracto y por elaborar modelos basados en la lógica deductiva[24]. Su postura lo enfrentó a Schmoller en la conocida **disputa del método** (*Methodenstreit*), que opuso a la escuela austriaca y al historicismo alemán a raíz de la obra de 1883 en la que Menger criticaba el método histórico.
  - *Escuela de Lausana*: L. Walras prescindió igualmente del tiempo y del lugar concretos, pero confiaba en que un uso profuso de las matemáticas permitiría entender la interdependencia y la causación mutua.

El resultado de esta época es el nacimiento del análisis de equilibrio general walrasiano (en su versión original), del análisis de equilibrio parcial marshalliano (que aísla mercados con la cláusula *ceteris paribus*) y del uso de la estadística para contrastar la teoría[25]. Por todo ello, puede decirse que en la Revolución Marginalista se encuentra el germen de la metodología económica moderna.

Revolución formalista moderna (1910-1970): a partir de 1910, la economía adopta los rasgos de una ciencia moderna, en cuatro procesos que se solapan en el tiempo:

▪ *Desarrollo de centros de investigación, revistas y econometría*: sobre el soporte metodológico de la Revolución Marginalista se configura una ortodoxia neoclásica que se mantiene más o menos inalterada hasta la década de 1930. En esta época se crean las instituciones de la investigación económica: departamentos universitarios, revistas (*American Economic Review*, 1911) y centros y sociedades de investigación (National Bureau of Economic Research, 1920; Sociedad Econométrica, 1930; Cowles Commission, 1932). De ellas nace la econometría. Tinbergen (1936, 1939) estima los primeros modelos macroeconométricos y Haavelmo (1944) asienta la disciplina en la teoría de la probabilidad. Sobre esa base, la Cowles Commission desarrolla los modelos de ecuaciones simultáneas, con los que se construyen después los grandes modelos macroeconométricos keynesianos.

▪ *Aparición del pensamiento keynesiano*: la Gran Depresión (1929-1933) y la obra de John Maynard Keynes *Teoría general de la ocupación, el interés y el dinero* (1936) provocaron una crisis del programa de investigación neoclásico, en especial en su vertiente marshalliana.

En este sentido, aunque Keynes no participó explícitamente en los debates metodológicos de la economía de los años treinta[26], la *Teoría general* (1936) respondía, en términos de Kuhn, a las

[24] Su método abstracto-deductivo lo acerca a Ricardo, aunque rechazó la teoría clásica del valor. Así, su obra prescinde de las matemáticas, de la estadística y del análisis de procesos históricos, si bien estudió el origen espontáneo de instituciones como el dinero.

[25] Jevons fue uno de los primeros en intentar realizar estudios empíricos macroeconómicos formales con la teoría de las manchas solares para entender los ciclos.

[26] Su padre, J. N. Keynes, sí había intervenido en la disputa del método, en 1891, con una posición conciliadora entre ambos enfoques.

anomalías del paradigma neoclásico. Este no podía explicar con sus propios fundamentos por qué las economías no recuperaban por sí solas el pleno empleo durante la Gran Depresión.

Así, si bien la contribución de Keynes no sustituyó el paradigma neoclásico[27], lo complementó y abrió la puerta al establecimiento de subobjetos de estudio (microeconomía y macroeconomía) con el fin de tratarlos con una profundidad que permitiese considerar realidades económicas como las imperfecciones de mercado en cada ámbito.

En el ámbito metodológico, el análisis keynesiano razona con relaciones entre agregados, como la función de consumo, que no se derivan de la decisión de cada agente, y de esa falta de fundamento individual partirá la crítica de Lucas (apartado 3.4.1). Cabe destacar también que la obra de Keynes *Treatise on Probability* (1921) es una contribución importante a la lógica de la inducción. Sin embargo, Keynes concibió la economía como ciencia moral, criticó el deductivismo de Ricardo[28] y desconfió de la econometría de su tiempo. En su reseña de Tinbergen (1939) objetó que la regresión exige un marco causal completo, factores medibles y relaciones estables, condiciones que la regresión no puede dar por supuestas.

▪ *Implantación del falsacionismo*: la crisis del paradigma neoclásico marshalliano coincide con un cambio en la norma metodológica de la disciplina. Frente al verificacionismo de Mill y Robbins, para el que una teoría vale por la certeza de sus premisas, cuatro autores trasladan entre 1938 y 1966 la prueba a las predicciones, que han de poder ser refutadas por los datos[29]. Así, Hutchison plantea la exigencia, Samuelson muestra cómo cumplirla, Friedman la convierte en el criterio para juzgar las teorías y Lipsey la lleva a los manuales.

- *Terence W. Hutchison*: el criterio de Popper suele considerarse introducido en la economía con su obra *The Significance and Basic Postulates of Economic Theory* (1938), aunque su exigencia de contrastabilidad se ha leído también como positivista (Caldwell, 1982) o como una posición propia de Hutchison (Hart, 2011). Hutchison introduce el criterio como crítica, porque no propone un método nuevo, sino que niega carácter empírico a buena parte de la teoría económica de su tiempo.

  Su blanco es el apriorismo de Robbins (1932), para quien los postulados de la teoría (que los individuos pueden ordenar sus preferencias, que los recursos son escasos) son hechos simples e indiscutibles de la experiencia que no necesitan contrastación. Hutchison distingue entre dos tipos de proposiciones, las tautológicas (imposibles de contrastar) y las empíricas (contrastables), y sostiene que la economía está dominada por las primeras, lo cual no supera el criterio de demarcación científico popperiano.

- *Paul Samuelson*: si Hutchison exige proposiciones contrastables, Samuelson muestra en su tesis doctoral, *Foundations of Economic Analysis* (1941, 1947), cómo obtenerlas de la teoría. Con ello, puede decirse, reintroduce la disciplina en el alcance del criterio de demarcación de Popper:

[27] La obra de Keynes también puede examinarse según la contextualización que hace Lakatos del proceso científico. Desde esta óptica, Keynes cuestionó el núcleo del programa neoclásico (la tendencia de una economía de mercado al pleno empleo mediante el ajuste de precios y salarios), más allá de reformular su cinturón protector. Por ello, sus ideas conformaron el núcleo de un nuevo programa de investigación alternativo al anterior. Este proceso continuó con las modificaciones del cinturón protector promovidas por los autores posteriores de la síntesis neoclásica y del monetarismo. Más tarde, la Nueva Macroeconomía Clásica implantó un nuevo programa de investigación cuyo núcleo difería del planteado por Keynes y por la síntesis neoclásica.

[28] Así, Keynes, en su ensayo sobre Malthus (1933), llega a afirmar que "la casi total desaparición del enfoque de Malthus y el dominio absoluto del de Ricardo durante cien años han sido un desastre para el progreso de la economía [...] ¡si Malthus, y no Ricardo, hubiera sido el tronco del que partió la economía del s. XIX, cuánto más sabio y rico sería hoy el mundo!".

[29] No todos estos autores mencionan explícitamente la obra de Karl Popper.

- *Teoremas operacionalmente significativos*: Samuelson fija como objetivo de la teoría obtener este tipo de teoremas, que define como hipótesis sobre datos empíricos que podrían concebiblemente refutarse, aunque fuera en condiciones ideales. El término procede del *operacionalismo* del físico Bridgman (1927), para quien un concepto se define por las operaciones con que se mide.
- *Principio de correspondencia*: las condiciones de equilibrio no bastan por sí solas para predecir en qué sentido cambia el equilibrio cuando cambia un parámetro (estática comparativa). Samuelson demuestra que el signo de ese cambio se obtiene de las condiciones de segundo orden, cuando el equilibrio procede de una optimización, y de la condición de estabilidad dinámica, cuando no procede de ella (de ahí el nombre del principio). Así, si un mercado es estable, un aumento de la demanda eleva el precio.

- *Milton Friedman*: con el ensayo *The Methodology of Positive Economics* (1953) se generaliza entre los economistas la idea de que una teoría se juzga solo por sus predicciones. Friedman coincide con Popper en que los hechos nunca pueden demostrar una hipótesis, sino solo no refutarla, y en que eso es lo que por lo general se quiere decir al afirmar que la experiencia la ha confirmado.

  En cambio, Friedman se aparta de Popper con la famosa y controvertida *tesis de irrelevancia de los supuestos*. Según esta tesis, que Samuelson (1963) denominó *F-Twist*, lo relevante no es el realismo de los supuestos, sino el poder predictivo de las conclusiones[30]. Se la suele calificar de *instrumentalista*, porque trata las teorías como instrumentos de predicción y no como descripciones verdaderas.

  Musgrave (1981) mostró que la tesis mezcla tres tipos de supuestos: de *irrelevancia* (un factor no influye apreciablemente), de *dominio* (delimitan el ámbito en que la teoría se aplica) y *heurísticos* (simplificaciones provisionales). El dictum de Friedman no vale para ninguno de los tres. Un supuesto de irrelevancia es una afirmación empírica que puede ser verdadera, uno de dominio falso deja la teoría sin aplicación y uno heurístico solo es aceptable como paso provisional.

- *Richard Lipsey*: su manual *An Introduction to Positive Economics* (1963) lleva el falsacionismo a la enseñanza. Tras el falsacionismo ingenuo de la primera edición, la segunda (1966) propone una falsación estadística, con un determinado nivel de confianza, y siempre comparada, porque una teoría solo se abandona cuando predice peor que una hipótesis alternativa. Con este manual, el falsacionismo pasa a ser la metodología declarada de la disciplina.

▪ *Extensión del formalismo matemático*: finalmente, cabe destacar que, durante esta etapa, se sistematiza y populariza el uso de las matemáticas, en microeconomía con la formulación axiomática y los programas de optimización y en macroeconomía con el modelo IS-LM (Hicks, 1937). La matematización se intensifica entre 1940 y 1970 con las obras de von Neumann y Morgenstern (1944), Samuelson (1947) y Arrow y Debreu (1954). Con Arrow y Debreu, el formalismo matemático se instaura en la metodología económica[31] y el enfoque walrasiano de equilibrio general desplaza en gran medida al marshalliano.

Microfundamentación y nuevas técnicas (1970-actualidad): desde los años setenta, la economía renueva sus métodos en tres campos:

▪ *Microeconomía*: al modelo estático de optimización individual y equilibrio competitivo se suma la *teoría de juegos*, que analiza las decisiones cuando el resultado de cada agente depende de

[30] Esta postura es diametralmente opuesta a la de Keynes, que, precisamente, critica a la escuela neoclásica por la falta de realismo de sus supuestos.

[31] Les precedieron autores que ya hacían un uso profuso de las matemáticas, como Jevons (1871), Edgeworth (1881) o Pareto (1906), relegados inicialmente por el predominio del análisis marshalliano.

lo que hagan los demás (tema A14) y que, desde los años ochenta, es el lenguaje común de la organización industrial (Tirole, 1988). A la vez, el experimento de laboratorio permite contrastar las predicciones de la teoría en condiciones controladas (apartado 3.1.1).

- *Macroeconomía*: los modelos pasan a ser dinámicos y estocásticos, es decir, describen decisiones a lo largo del tiempo sujetas a perturbaciones aleatorias. Sus técnicas, el *control óptimo* y la *programación dinámica*, entran en la economía en los años sesenta con la teoría del crecimiento óptimo (Cass, 1965; Koopmans, 1965; tema A29). La crítica de Lucas (1976) las generaliza después a toda la macroeconomía.

- *Econometría*: la práctica econométrica de la posguerra recibe tres críticas, que abren otros tantos modos de trabajar. De la de Lucas (1976) procede la estimación de modelos estructurales, cuyos parámetros son las preferencias y la tecnología; de la de Sims (1980), los modelos VAR, y de la de Leamer (1983), los diseños cuasiexperimentales, que buscan identificar efectos causales (apartados 3.4 y 3.1.3). A ellos se suman el análisis de series temporales no estacionarias, la microeconometría y la estadística bayesiana (Zellner, 1971).

2.2.3. Concepto actual

Características: la economía comparte con las demás ciencias empíricas la exigencia de contrastación, pero su método presenta dos rasgos propios. Por una parte, sus leyes son *inexactas*, porque describen tendencias que solo se cumplen *ceteris paribus*, y su contrastación es difícil, porque los datos proceden casi siempre de la observación y no del experimento (apartado 3.1). Por otra, la economía explica su ámbito con un número reducido de factores causales, deducidos de supuestos sobre la conducta individual. Hausman (1992) la describe por ello como una ciencia *inexacta y separada*, heredera del método *a priori* de Mill.

La contrastación empírica y el formalismo han hecho predominar en la economía actual un método formal y cuantitativo, con el que conviven otras metodologías. Según Gabaix y Laibson (*The Seven Properties of Good Models*, 2008), un buen modelo económico reúne siete propiedades, que pueden agruparse en tres requisitos:

- *Formalización*: profundidad conceptual, capacidad de generalización y parsimonia, es decir, explicar mucho con pocos supuestos.

- *Aplicabilidad*: los modelos han de ser falsables (contrastables con los datos) y tratables (resolubles con las herramientas disponibles).

- *Utilidad*: consistencia empírica y precisión predictiva.

Estas propiedades no siempre son compatibles, y por eso elegir un modelo exige ponderarlas según la pregunta que se quiere responder.

Debates: en el plano metodológico siguen abiertas cuatro discusiones:

1. *Monismo y pluralismo*: se discute en qué medida difiere el método económico del de otras ciencias. Los partidarios del monismo metodológico sostienen que el distinto objeto de las ciencias no debería influir en su método, mientras que los del pluralismo metodológico sostienen que las diferencias entre las ciencias van más allá de su objeto[32].

[32] Esta posición se inspira en la noción de *Verstehen* de la filosofía alemana del s. XIX (comprensión desde dentro, por intuición, según Blaug, 1992). A diferencia del científico natural, que no puede "sentir ser" un átomo, el científico social puede y debe comprender a los individuos cuyo comportamiento estudia, por lo que no puede limitarse a la observación externa.

2. *Falsacionismo en la práctica*: Blaug (1992) concluye que los economistas suscriben el falsacionismo, pero rara vez lo practican, y llama a esa actitud *falsacionismo inocuo* (*innocuous falsificationism*). También Hausman (1992) critica el dogmatismo con que la disciplina, fiel al método de Mill, protege sus leyes básicas frente a la evidencia contraria. Caldwell (1982), por su parte, concluye que ninguna metodología ha demostrado su superioridad y propone un *pluralismo metodológico* crítico. El término alude aquí a la convivencia de metodologías evaluadas con criterios explícitos, no a la diferencia entre ciencias.

3. *Individualismo metodológico y holismo*: el *individualismo metodológico* sostiene que los fenómenos sociales deben explicarse a partir de las decisiones e interacciones de los individuos. La expresión procede de Schumpeter (1908), y la idea, de Weber y Menger. El *holismo*, en cambio, atribuye a las clases, las instituciones o la sociedad regularidades propias, irreducibles a los individuos (Durkheim, Marx). La economía dominante es individualista, como muestra la microfundamentación de la macroeconomía, aunque el agente representativo es una agregación que el individualismo estricto no justifica (Kirman, 1992) y la economía de la complejidad estudia propiedades que emergen de la interacción (apartado 3.3).

4. *Formalismo y retórica*: McCloskey (1985) sostiene que los economistas persuaden mediante recursos retóricos (metáforas, autoridad, el prestigio de las matemáticas) más que mediante la contrastación que proclama su metodología oficial. Por su parte, P. Romer (2015) denomina *mathiness* al uso de las matemáticas que deja holgada la conexión entre los símbolos y los conceptos económicos, y entre la teoría y los datos.

2.3. Conclusiones

El método científico no es un conjunto fijo de reglas, y la reflexión sobre él, del deductivismo a la nueva filosofía de la ciencia, ha girado en torno a qué cuenta como contrastación suficiente.

La economía ha recorrido un camino paralelo, del verificacionismo de Mill y Robbins a la exigencia de predicciones refutables. En la práctica, sin embargo, se aleja a menudo del falsacionismo que proclama, como subrayan Blaug y Hausman. Con todo, su método combina hoy la abstracción teórica y la contrastación empírica en un ciclo que va de la realidad a los modelos y de estos a los datos.

3. Cuestiones y debates actuales

Junto a los debates metodológicos del apartado 2.2.3, desde la segunda mitad del s. XX la economía experimental, la economía conductual y la economía de la complejidad discuten dos supuestos tradicionales de la disciplina, el de que la economía no puede experimentar y el de que sus agentes son plenamente racionales y se encuentran en equilibrio. A estas tres líneas de investigación se suma el debate sobre el método de la macroeconomía, de la crítica de Lucas a la crisis de 2008.

Las tres tienen en común que someten a contraste el núcleo neoclásico y lo amplían, sin sustituirlo. El reconocimiento académico de las dos primeras se refleja en los Premios Nobel de 2002, 2017, 2019 y 2021.

3.1. Economía experimental (1960-actualidad)

La **economía experimental** discute, con experimentos de laboratorio y de campo, la tesis tradicional de que la economía no puede realizar experimentos controlados. En el laboratorio, el investigador crea un entorno económico artificial, y en el trabajo de campo estudia el comportamiento en entornos reales mediante asignaciones aleatorias, sean estas diseñadas por el investigador o fruto de las circunstancias.

3.1.1. Experimentos de laboratorio: Vernon Smith (1962)

Origen: Vernon Smith (1962) es considerado el padre de la economía experimental. Recibió el Premio Nobel en 2002 por haber establecido los experimentos de laboratorio como herramienta de análisis económico empírico, en especial para el estudio de mecanismos alternativos de mercado.

Idea: el investigador reproduce en el laboratorio una situación de mercado y controla las variables que en la realidad no se observan. Smith asignaba a cada participante un valor de reserva (precio máximo para los compradores, coste para los vendedores), de modo que las curvas de oferta y demanda inducidas eran conocidas de antemano. Después, Smith dejaba negociar a los participantes mediante una *subasta doble*, en la que compradores y vendedores anuncian públicamente pujas y ofertas.

Razonamiento: su aportación es doble:

- *Resultado*: con pocos participantes y sin que ninguno conociera las curvas, los precios y las cantidades convergían con rapidez al equilibrio competitivo, lo que sugiere que esa convergencia no exige información perfecta ni un gran número de agentes.

- *Método*: Smith fijó después las reglas del diseño experimental, entre ellas la *teoría del valor inducido* (Smith, 1976), según la cual, si la recompensa monetaria del participante depende de sus decisiones, el investigador controla sus preferencias y puede contrastar predicciones teóricas precisas.

Extensiones: la generación siguiente profundizó en el método, y Alvin Roth (Nobel 2012, con Shapley) combinó la teoría de emparejamientos, el laboratorio y el estudio de campo en el diseño de mercados reales.

Conclusiones:

- *Aspectos positivos*: el laboratorio ofrece *validez interna* (control de las condiciones).

▪ *Limitaciones*: su *validez externa* es discutida, porque los participantes suelen ser estudiantes, las cantidades en juego son pequeñas y el entorno es artificial.

3.1.2. Ensayos controlados aleatorios: Banerjee, Duflo y Kremer (Nobel 2019)

Origen: el campo principal de los ensayos controlados aleatorios es la economía del desarrollo, donde Banerjee, Duflo y Kremer (Nobel 2019) descompusieron los grandes interrogantes de la pobreza en preguntas acotadas y comprobables (incentivos a profesores, subsidios sanitarios, microcréditos).

Idea: en un *ensayo controlado aleatorio* (*randomised controlled trial*, RCT) el investigador controla las condiciones del experimento en el terreno.

Razonamiento:

▪ *Asignación aleatoria*: los participantes se asignan al azar, por decisión del investigador, a un grupo de tratamiento (que recibe la intervención) o a un grupo de control (que no la recibe). El azar garantiza que ambos grupos sean, en promedio, iguales en todo salvo en el tratamiento.

▪ *Efecto causal*: por ello, la diferencia de las medias muestrales entre ambos grupos es un estimador insesgado del efecto causal medio del tratamiento:

$$\widehat{ATE} = \bar{Y}_{T=1} - \bar{Y}_{T=0}, \qquad ATE = E\big[Y_i(1) - Y_i(0)\big]$$

donde $Y_i(1)$ e $Y_i(0)$ son los resultados potenciales del individuo i (renta, asistencia escolar, empleo) con y sin tratamiento, de los que solo se observa uno. Las medias muestrales del resultado en los grupos de tratamiento y de control son $\bar{Y}_{T=1}$ y $\bar{Y}_{T=0}$, y $\widehat{ATE}$ es la estimación del efecto medio del tratamiento (*average treatment effect*).

Extensiones: antes de su auge en el desarrollo, los ensayos controlados aleatorios se usaron en EE. UU. desde 1968 para evaluar programas sociales (impuesto negativo sobre la renta, seguro sanitario de RAND), y hoy evalúan también políticas activas de empleo.

> La aproximación experimental a la economía del desarrollo, sus principales resultados y sus críticas (validez externa y escalabilidad) se desarrollan en el tema B28, apartado 3.3.

Conclusiones:

▪ *Aspectos positivos*: la asignación aleatoria elimina el sesgo de selección y permite estimar efectos causales con supuestos mínimos. Por ello, los ensayos han orientado la evaluación de políticas públicas y la economía del desarrollo hacia la evidencia medible.

▪ *Limitaciones*: al igual que en el laboratorio, su validez externa es discutida, pues el efecto medido en una población y un contexto concretos puede no repetirse en otros ni a mayor escala. Además, son costosos, plantean dilemas éticos y solo responden a preguntas que admiten aleatorización.

3.1.3. Experimentos naturales: Card, Angrist e Imbens (Nobel 2021)

Origen: David Card compartió el Premio Nobel de 2021 con Joshua Angrist y Guido Imbens. Card lo recibió por sus contribuciones empíricas a la economía laboral, y Angrist e Imbens, por sus aportaciones metodológicas al análisis de relaciones causales.

Idea: en un *experimento natural* el investigador no controla la asignación. Un factor ajeno, natural o institucional, expone a unos individuos a un tratamiento y a otros no (una frontera, una catástrofe natural, un cambio normativo).

Razonamiento: la tarea del investigador es doble:

- *Comparabilidad*: comprobar que el grupo de tratamiento y el de control son similares en lo observable.

- *Plausibilidad*: argumentar que el factor que determina la exposición puede tratarse *como si* fuera una asignación aleatoria, es decir, que no está correlacionado con otros determinantes del resultado.

Angrist e Imbens mostraron qué conclusiones causales precisas pueden extraerse de un experimento natural cuando los individuos deciden en parte si reciben o no el tratamiento. En concreto, puede estimarse el efecto medio local del tratamiento (LATE, Imbens y Angrist, 1994), es decir, el efecto sobre quienes reciben el tratamiento solo porque el instrumento (p. ej. un sorteo) los empuja a ello, los llamados *cumplidores*. Entre los resultados de Card destacan:

1. *Salario mínimo* (Card y Krueger, 1994): la subida del salario mínimo en Nueva Jersey, comparada con la vecina Pensilvania, no redujo el empleo en los restaurantes de comida rápida.

2. *Inmigración* (Card, 1990): la llegada masiva de inmigrantes a Miami tras el éxodo del Mariel (1980) no tuvo efectos apreciables sobre los salarios ni sobre el empleo de los trabajadores locales poco cualificados, resultado discutido después por Borjas (2017).

Conclusiones:

- *Aspectos positivos*: los experimentos naturales y las técnicas asociadas (variables instrumentales, diferencias en diferencias, regresión discontinua) han dado lugar a la llamada *revolución de la credibilidad* en la economía empírica (Angrist y Pischke, 2010). Esta revolución responde a la crítica de Leamer (1983) a una econometría cuyos resultados dependían de especificaciones elegidas por el investigador. La identificación de efectos causales a partir de una variación creíblemente exógena ha pasado a ser el estándar de la disciplina.

- *Limitaciones*: sus críticos (Deaton, 2010; Heckman, 2010) señalan que estos métodos identifican efectos locales, en una población y un contexto concretos, que no siempre pueden extrapolarse, y que sin teoría económica no explican por qué se produce el efecto.

3.2. Economía conductual (1970-actualidad)

La **economía conductual** (*behavioral economics*) incorpora a la economía los hallazgos de la psicología sobre cómo deciden realmente las personas. Su principal antecedente es la **racionalidad limitada** de Herbert Simon (1955, 1957; Nobel 1978), aunque la economía conductual actual, desde Kahneman y Tversky, conserva la optimización como referencia y mide las desviaciones respecto de ella (Sent, 2004). Para Simon, los agentes tienen una capacidad limitada para procesar información, por lo que, en lugar de optimizar, se conforman con alternativas satisfactorias. Frente al *homo economicus*, la economía conductual documenta desviaciones sistemáticas y

predecibles respecto de la racionalidad, sesgos con una dirección definida y no errores aleatorios que se compensan en el agregado. Las preferencias sociales (juego del ultimátum, aversión a la desigualdad) se recogen en el Anexo A.3.

3.2.1. El descuento hiperbólico (Herrnstein, 1961; Ainslie, 1975)

Origen: Richard Herrnstein (1961) formuló, en experimentos con animales, la *ley de igualación*, que establece que la frecuencia relativa de una conducta se iguala a la frecuencia relativa de la recompensa que obtiene. De ella se deriva que el valor de una recompensa diferida cae de forma hiperbólica con la demora, idea que George Ainslie (1975) aplicó a la elección humana.

Idea: en el modelo estándar de Samuelson (1937), el descuento es exponencial y la tasa de descuento entre dos periodos no depende de cuándo se evalúe, por lo que las preferencias son coherentes en el tiempo. En cambio, con descuento hiperbólico, la tasa es alta a corto plazo y baja a largo, y las preferencias son *dinámicamente inconsistentes*, problema que Strotz (1956) había analizado ya en economía. Así, un individuo prefiere hoy 110 euros dentro de 31 días a 100 euros dentro de 30, pero, llegado el día 30, prefiere los 100 inmediatos.

Razonamiento: David Laibson (1997) formalizó esta conducta con el descuento cuasihiperbólico (propuesto por Phelps y Pollak, 1968), que explica fenómenos como el ahorro insuficiente para la jubilación o la demanda de mecanismos de compromiso, como los planes de pensiones con penalización por rescate anticipado.

Conclusiones:

▪ *Aspectos positivos*: un único sesgo hacia el presente explica la inconsistencia dinámica y la demanda de mecanismos de compromiso, y la versión cuasihiperbólica, con un solo parámetro más que la exponencial, es tratable.

▪ *Limitaciones*: sus predicciones dependen de que el individuo sea *sofisticado* (conoce su sesgo) o *ingenuo* (O'Donoghue y Rabin, 1999), algo difícil de observar. Además, en los experimentos con dinero la preferencia por el pago inmediato puede reflejar desconfianza en el pago futuro más que descuento.

3.2.2. Teoría de la perspectiva y heurísticas (Tversky y Kahneman, 1974; Kahneman y Tversky, 1979)

Origen: Daniel Kahneman (Nobel 2002) y Amos Tversky, psicólogos, integraron en la economía la investigación sobre el juicio y la toma de decisiones.

Idea: los individuos juzgan y eligen en condiciones de incertidumbre con atajos y valoraciones que se apartan de la utilidad esperada y de las leyes de la probabilidad de forma sistemática y predecible.

Razonamiento: su obra se articula en tres bloques:

▪ *Teoría de la perspectiva* (Kahneman y Tversky, 1979): sustituye la utilidad esperada por una función de valor definida sobre ganancias y pérdidas y una función de ponderación de las probabilidades. Sus cuatro elementos son:

- *Dependencia del punto de referencia*: los agentes valoran cambios respecto a un punto de referencia, no niveles absolutos de riqueza.
- *Aversión a la pérdida*: una pérdida pesa más que una ganancia de igual magnitud.

- *Sensibilidad decreciente*: la función de valor es cóncava en ganancias y convexa en pérdidas (véase la figura del tema A10, apartado 2.3.2).

- *Ponderación de probabilidades*: se sobreponderan las probabilidades bajas y se infraponderan las moderadas y altas.

▪ *Reglas heurísticas* (Tversky y Kahneman, 1974): estos autores identifican los atajos cognitivos que, aunque útiles, llevan a los individuos a cometer errores sistemáticos:

- *Representatividad*: se juzga la probabilidad de un suceso por su parecido con un estereotipo, ignorando la probabilidad de partida.

- *Disponibilidad*: se juzga la frecuencia de un suceso por la facilidad con que vienen a la mente ejemplos de él.

- *Anclaje y ajuste*: se parte de un valor inicial (el ancla) y el ajuste posterior es insuficiente.

▪ *Efectos*: derivados de lo anterior, contradicen la teoría estándar:

- *Efecto marco* (Tversky y Kahneman, 1981): presentar una misma decisión con connotaciones positivas o negativas (p. ej. "salvar 200 vidas de 600" frente a "dejar morir a 400 de 600") cambia la elección, en contra del principio de invarianza.

- *Efecto justicia* (Kahneman, Knetsch y Thaler, 1986): las decisiones económicas se ven influidas por lo que se considera justo (p. ej. el rechazo a subir el precio de las palas quitanieves tras una gran nevada).

> La teoría de la perspectiva se desarrolla en el tema A10, apartado 2.3, y su versión acumulativa (Tversky y Kahneman, 1992), en el apartado 3.1.2.

Conclusiones:

▪ *Aspectos positivos*: la teoría de la perspectiva explica, con supuestos contrastados en el laboratorio, elecciones que la utilidad esperada no recoge, como la aversión al riesgo en las ganancias y su búsqueda en las pérdidas.

▪ *Limitaciones*: la teoría no dice cómo fijar el punto de referencia, lo que resta capacidad predictiva. Además, su versión de 1979 puede violar la dominancia estocástica de primer orden, defecto que corrige la versión acumulativa.

3.2.3. Contabilidad mental, efecto dotación y *nudge* (Thaler, 1980; 1985; Thaler y Sunstein, 2008)

Origen: Richard Thaler (Nobel 2017) recibió el premio por incorporar supuestos psicológicamente realistas al análisis de las decisiones económicas.

Idea: la racionalidad limitada, las preferencias sociales y la falta de autocontrol afectan de forma sistemática a las decisiones individuales y a los resultados de los mercados.

Razonamiento: sus aportaciones son cuatro:

- *Contabilidad mental* (Thaler, 1985): los individuos clasifican el dinero en cuentas mentales separadas (gastos corrientes, ocio, ahorro) y no lo tratan como fungible, en contra del supuesto estándar. Así, como ya habían mostrado Tversky y Kahneman (1981), quien pierde una entrada de teatro es menos propenso a comprar otra que quien pierde el equivalente en efectivo (46 % frente a 88 %).

- *Teoría del "empujoncito"* (*nudge*; Thaler y Sunstein, 2008): la forma en que se presentan las opciones influye en la decisión sin restringirlas ni alterar sus incentivos. Quienes diseñan ese contexto se convierten en *arquitectos de la elección*. Por ejemplo, la inscripción automática en planes de pensiones de empresa, con posibilidad de salir, eleva notablemente la participación sin restringir la libertad de elección (el llamado *paternalismo libertario*).

- *Efectos*: destaca el *efecto dotación* o de propiedad (Thaler, 1980), por el que el individuo valora más un bien cuando lo posee que cuando no lo posee, lo que abre una brecha entre la disposición a pagar y la disposición a aceptar. Thaler (1985) añade la *utilidad de transacción*, el valor que se atribuye a la calidad del trato.

- *Autocontrol* (Thaler y Shefrin, 1981): los agentes tienen dificultades para cumplir sus propios planes. Thaler y Shefrin modelizan al individuo como un *planificador* con visión de largo plazo y un *ejecutor* miope, y explican por qué se recurre a reglas y compromisos (ahorro obligatorio, domiciliaciones).

> La racionalidad limitada, el "empujoncito" y el efecto dotación se aplican a la función de consumo en el tema A33, apartado 3.2.3.

Conclusiones:

- *Aspectos positivos*: su obra da explicaciones sencillas y contrastables de conductas cotidianas de consumo y ahorro que el modelo estándar trata como anomalías.

- *Limitaciones*: la brecha entre las disposiciones a aceptar y a pagar desaparece en experimentos que controlan los errores de comprensión de los participantes (Plott y Zeiler, 2005), lo que discute el efecto dotación. Además, el *nudge* se critica por influir en la decisión sin que el individuo lo advierta.

3.2.4. Conclusiones

Aspectos positivos: la economía conductual ha dado lugar a políticas públicas basadas en el comportamiento, desde el Behavioural Insights Team británico (2010) a unidades similares en otros Gobiernos y organismos.

Limitaciones: su éxito ha abierto cuatro debates:

- *Teoría unificada*: la economía conductual acumula sesgos sin una teoría que prediga cuál domina en cada contexto. Gigerenzer añade que muchas heurísticas son reglas eficientes con información escasa (*racionalidad ecológica*) y no errores.

- *Replicación*: de 18 experimentos de laboratorio publicados en la *American Economic Review* y el *Quarterly Journal of Economics* entre 2011 y 2014, 11 (el 61 %) se replicaron (Camerer et al., 2016).

- *Escala*: en dos grandes unidades de comportamiento de EE. UU., un *nudge* elevó la respuesta en 1,4 puntos porcentuales de media, frente a 8,7 en los estudios publicados en revistas, sobre todo por el sesgo de publicación (DellaVigna y Linos, 2022).

- *Bienestar*: si las elecciones no reflejan el bienestar, hace falta otro criterio para evaluar políticas, y el regulador que corrige sesgos también puede tenerlos.

3.3. Economía de la complejidad (1987-actualidad)

La **economía de la complejidad** estudia la economía como un **sistema complejo adaptativo**, formado por muchos agentes heterogéneos que interactúan, aprenden y se adaptan, y cuyo comportamiento agregado surge de esas interacciones sin que nadie lo diseñe. Frente al análisis de equilibrio, se interesa por los procesos, es decir, por cómo se forman las pautas agregadas y cómo cambian.

3.3.1. El programa del Instituto Santa Fe: el mercado bursátil artificial (Arthur et al., 1997)

Origen: en septiembre de 1987 el Instituto Santa Fe (Nuevo México) reunió, por iniciativa de Kenneth Arrow y Philip Anderson, a economistas y a científicos de la física, la biología y la informática. El encuentro tenía por objeto estudiar la economía como un sistema complejo en evolución (Anderson, Arrow y Pines, 1988). En 1988 se creó su programa de economía, dirigido por W. Brian Arthur.

Idea: en ese marco, W. Brian Arthur y John Holland, con otros investigadores, construyeron a finales de los años ochenta un *mercado bursátil artificial* con una "ecología" de estrategias de negociación. Este mercado es el ejemplo de referencia de la economía de la complejidad (Arthur, Holland, LeBaron, Palmer y Tayler, 1997).

Razonamiento: en este mercado, agentes heterogéneos forman sus expectativas mediante reglas de predicción que ponen a prueba y sustituyen según su éxito, en lugar de suponer expectativas racionales. Cuando el aprendizaje es lento, los precios convergen al equilibrio de expectativas racionales, mientras que, cuando es rápido, el mercado genera por sí solo burbujas, crisis, volatilidad agrupada y un volumen de negociación elevado.

El mercado artificial mostró así que el equilibrio de expectativas racionales es solo uno de los regímenes posibles y dio impulso a los *modelos basados en agentes*, hoy utilizados, entre otros ámbitos, en el análisis de la estabilidad financiera.

Conclusiones:

- *Aspectos positivos*: ofrece una alternativa al agente representativo con expectativas racionales y explica desde la propia dinámica del mercado regularidades empíricas que el modelo estándar no recoge.

- *Limitaciones*: sus resultados dependen de reglas de comportamiento y de parámetros elegidos por el investigador, con muchos grados de libertad y difíciles de contrastar con los datos. Además, la simulación no ofrece resultados analíticos generales, lo que dificulta comparar modelos y saber qué mecanismo produce cada resultado.

3.3.2. Propiedades de los sistemas complejos (Arthur, Durlauf y Lane, 1997)

Origen: W. Brian Arthur, Steven Durlauf y David Lane sintetizaron en 1997, en la introducción a *The Economy as an Evolving Complex System II*, la perspectiva desarrollada en el Instituto Santa Fe durante su primera década.

Idea: la economía es un sistema complejo adaptativo, que se reconoce por seis propiedades que el análisis de equilibrio no recoge.

Razonamiento: las seis propiedades son las siguientes:

- *Interacción dispersa*: la actividad económica resulta de la interacción de muchos agentes heterogéneos que actúan en paralelo.

- *Ausencia de un controlador global*: no hay una autoridad que coordine todas las interacciones, y el control lo ejercen los mecanismos de competencia y coordinación entre agentes.

- *Organización jerárquica transversal*: la economía se organiza en niveles (unidades, organizaciones, sectores), y cada nivel sirve de base al siguiente, con múltiples relaciones entre ellos.

- *Adaptación continua*: los agentes revisan sus conductas y estrategias a medida que acumulan experiencia.

- *Novedad perpetua*: surgen continuamente nuevos mercados, tecnologías, instituciones y conductas, que abren nuevos nichos.

- *Dinámica fuera del equilibrio*: por todo lo anterior, la economía opera lejos de cualquier óptimo o equilibrio global, y las mejoras son siempre posibles.

Conclusiones:

- *Aspectos positivos*: ofrece un marco común para modelos muy distintos (mercados artificiales, redes, modelos basados en agentes) y explica por qué la economía genera novedad y cambio estructural de forma continua.

- *Limitaciones*: las propiedades describen el sistema más que predecir su comportamiento, y su traducción en hipótesis contrastables depende de modelos concretos, con los problemas de calibración y validación ya señalados.

3.3.3. El índice de complejidad económica (Hidalgo y Hausmann, 2009)

Origen: César Hidalgo y Ricardo Hausmann (2009) aplicaron la perspectiva de la complejidad al desarrollo económico.

Idea: la riqueza de un país depende de la cantidad y diversidad de capacidades productivas que acumula (conocimientos, instituciones, infraestructuras). Esas capacidades no se observan directamente, pero se reflejan en lo que el país exporta.

Razonamiento: el *índice de complejidad económica* (ICE), formalizado en el *Atlas de Complejidad Económica* (Hausmann, Hidalgo et al., 2011), combina dos dimensiones:

- *Diversidad*: cuántos productos exporta un país con ventaja comparativa revelada.

- *Ubicuidad*: cuántos países exportan cada uno de esos productos. Un país es complejo si exporta muchos productos que pocos países son capaces de producir.

Según el *Atlas*, el índice predice el crecimiento futuro mejor que indicadores convencionales de gobernanza, competitividad o capital humano. Los países cuya complejidad es alta en relación con su renta tienden a crecer más deprisa.

Extensiones: el Growth Lab de la Universidad de Harvard publica el índice, que se utiliza para identificar las diversificaciones productivas más accesibles a partir de las capacidades existentes.

Conclusiones:

- *Aspectos positivos*: resume en un solo indicador, calculable con datos de comercio disponibles para casi todos los países, esas capacidades productivas inobservables.

- *Limitaciones*: en su formulación original descansa solo en las exportaciones de bienes, por lo que omite los servicios y la producción para el mercado interior, y en las cadenas globales de valor exportar un bien complejo puede reflejar solo su ensamblaje. Además, su capacidad predictiva es una asociación estadística que no prueba por sí sola una relación causal.

3.4. El método de la macroeconomía (1976-actualidad)

El debate sobre el método de la macroeconomía arranca de las críticas de Lucas y de Sims a los grandes modelos macroeconométricos y se reabre con la crisis de 2008.

3.4.1. La crítica de Lucas (1976)

Origen: ante el fracaso de la curva de Phillips en la estanflación de los años setenta, Robert Lucas (Nobel 1995) cuestiona en *Econometric Policy Evaluation: A Critique* (1976) la práctica, entonces habitual, de evaluar las políticas simulándolas en los grandes modelos macroeconométricos (apartado 2.2.2).

Idea: las ecuaciones de esos modelos recogen las reglas de decisión de los agentes, y estas reglas dependen de lo que los agentes esperan de la política económica. Por tanto, cuando la política cambia, cambian las reglas y, con ellas, los coeficientes estimados.

Razonamiento: de esta idea se siguen dos consecuencias de método:

- *Invalidez de la simulación de políticas*: una relación estimada bajo un régimen de política deja de cumplirse cuando se adopta otro. Así, la curva de Phillips estimada con inflación baja y estable se desplazó cuando los Gobiernos intentaron explotarla, porque los agentes revisaron sus expectativas.

- *Microfundamentación*: solo sirven para evaluar políticas los modelos cuyos parámetros no dependen de la política, es decir, las preferencias y la tecnología (*parámetros profundos*). Ello obliga a derivar las relaciones agregadas de la optimización de agentes con expectativas racionales (tema A06, apartados 1.2 y 1.3).

Extensiones: de esta exigencia nacen los modelos de ciclo real (Kydland y Prescott, 1982), que se *calibran* con evidencia microeconómica en lugar de estimarse, y, después, los *modelos dinámicos estocásticos de equilibrio general* (DSGE; tema A06, apartado 2.3).

Conclusiones:

- *Aspectos positivos*: la crítica cambió el modo de hacer macroeconomía, y suele describirse como un cambio de paradigma en el sentido de Kuhn (apartado 2.1.2).

- *Limitaciones*: su relevancia empírica ha resultado menor de lo que se temía (tema A06, apartado 1.3). Además, la exigencia de microfundamentos llevó a relegar los modelos que no la cumplían, lo que se reprochó a la disciplina tras 2008.

3.4.2. La crítica de Sims y los modelos VAR (1980)

Origen: Christopher Sims (Nobel 2011) dirige en *Macroeconomics and Reality* (1980) una segunda crítica a los grandes modelos macroeconométricos, que no se refiere a su uso, como la de Lucas, sino a su identificación.

Idea: para estimar un modelo de ecuaciones simultáneas hay que *identificarlo*, es decir, imponer restricciones que permitan separar unas relaciones de otras. Los grandes modelos lo hacían excluyendo variables de cada ecuación y declarando exógenas otras, sin que la teoría lo justificase, y Sims califica por ello sus restricciones de "increíbles".

Razonamiento: como alternativa, Sims propone los *modelos de vectores autorregresivos* (VAR). En ellos, todas las variables se tratan como endógenas y cada una se explica por sus propios retardos y por los de las demás, sin restricciones tomadas de la teoría. El modelo se interpreta con las *funciones de impulso-respuesta*, que describen cómo reacciona cada variable a lo largo del tiempo ante una perturbación inesperada en otra.

Extensiones: interpretar una perturbación como un *shock* de política exige, sin embargo, algún supuesto de identificación. Por ello, desde mediados de los años ochenta se desarrollan los *VAR estructurales* (SVAR), con los que se estiman, por ejemplo, los multiplicadores fiscal y monetario (tema A40, apartado 3.1).

Conclusiones:

- *Aspectos positivos*: los VAR dejan que los datos muestren la dinámica de la economía y son hoy una herramienta habitual de predicción y de análisis de políticas.

- *Limitaciones*: Cooley y LeRoy (1985) los calificaron de *ateóricos*, porque sin restricciones teóricas no permiten extraer conclusiones causales. Además, sus coeficientes son de forma reducida, de modo que tampoco escapan a la crítica de Lucas.

3.4.3. La macroeconomía tras la crisis de 2008

La Gran Recesión reabrió el debate metodológico en macroeconomía, porque los modelos DSGE dominantes no la anticiparon, en su mayoría carecían de fricciones financieras (pese a precedentes como Bernanke, Gertler y Gilchrist, 1999), descansaban en un agente representativo y trataban mal el límite inferior de los tipos de interés (*zero lower bound*, ZLB). Romer (*The Trouble with Macroeconomics*, 2016) llegó a sostener que la disciplina había retrocedido al anteponer la coherencia formal a la evidencia.

La respuesta mayoritaria ha sido ampliar el marco antes que abandonarlo, con fricciones financieras y agentes heterogéneos (modelos *Heterogeneous Agent New Keynesian*, HANK, de Kaplan, Moll y Violante, 2018), y aceptar un pluralismo de herramientas. Blanchard (2018), por ejemplo, distingue entre modelos fundacionales, DSGE, de política, de juguete y de predicción, cada uno con su función. A ello se suma el uso creciente de datos masivos y del aprendizaje automático, útil para predecir, pero que no sustituye la identificación causal (Mullainathan y Spiess, 2017).

Los límites de los modelos DSGE ante la Gran Recesión y las respuestas posteriores (fricciones financieras, agentes heterogéneos) se desarrollan en el tema A06, apartado 2.3. Aquí se recoge su dimensión metodológica.

Conclusiones

El tema ha delimitado primero el objeto de la economía. Su definición ha pasado del arte de gobierno de Smith y la ciencia de la riqueza de Say, Mill y Marshall a la definición analítica de Robbins, la elección entre fines y medios escasos con usos alternativos. La de Robbins es hoy la más aceptada por su generalidad, aunque esa misma generalidad la expone a críticas. En cuanto al método, el procedimiento hipotético-deductivo combina abducción, deducción e inducción, y la reflexión sobre él ha pasado del verificacionismo a la falsación de Popper y, con Kuhn, Lakatos y Feyerabend, al estudio de cómo evoluciona de hecho la ciencia. La economía ha seguido un camino paralelo, del método *a priori* de Mill a la contrastación econométrica, aunque su práctica se aleja a menudo del falsacionismo que proclama. Por último, la economía experimental, la conductual y la de la complejidad discuten que la economía no pueda experimentar y que sus agentes sean plenamente racionales.

Estas cuestiones de objeto y de método recorren el resto del libro. La deducción a partir de supuestos introspectivos, que Mill defendió frente a la inducción histórica, es el método de la escuela clásica (tema A02). La exigencia de que un modelo sirva para evaluar políticas, y no solo para ajustar los datos, está detrás de la crítica de Lucas, según la cual los parámetros estimados cambian con la propia política (tema A06). Asimismo, la racionalidad limitada y los sesgos que la economía conductual opone al agente racional reaparecen en la elección en condiciones de incertidumbre (tema A10). En todos los casos, el modelo simplificado, contrastable y abierto a la refutación es el método que siguen los demás temas.

La revolución de la credibilidad (Angrist y Pischke, 2010), respuesta a la crítica de Leamer (1983), ha hecho de los diseños cuasiexperimentales el estándar de la economía empírica, aunque identifican efectos locales difíciles de extrapolar sin modelos estructurales (apartado 3.1.3). Sigue abierta, además, la fiabilidad de la evidencia publicada. Camerer et al. (2016) solo lograron replicar 11 de 18 experimentos de laboratorio publicados en dos grandes revistas, y el sesgo de publicación agrava el problema (apartados 2.1.3 y 3.2.4). En macroeconomía, por último, Blanchard (2018) defiende la convivencia de varias clases de modelos tras la Gran Recesión, y Mullainathan y Spiess (2017) advierten de que el aprendizaje automático predice bien, pero no sustituye la identificación causal (apartado 3.4).

Anexos

A.1. Corrientes posteriores de la filosofía de la ciencia

Desde los años setenta, otras tres líneas completan el panorama del apartado 2.1.2. La *sociología del conocimiento científico*, con el programa fuerte de Bloor (1976), explica la aceptación de las teorías por los intereses y las normas de las comunidades científicas, y no solo por la evidencia. El análisis *retórico* estudia cómo persuaden los científicos y McCloskey lo aplica a la economía (apartado 2.2.3). Por último, el *realismo científico* sostiene, frente al instrumentalismo, que las teorías que tienen éxito describen, al menos de forma aproximada, estructuras reales del mundo. En economía, Lawson (1997) desarrolla, a partir de Bhaskar (1975), un *realismo crítico* según el cual la realidad social es un sistema abierto de estructuras y mecanismos subyacentes, lo que le lleva a rechazar el deductivismo de la economía dominante.

A.2. Metodología de la escuela austriaca

Siguiendo a Ludwig von Mises (*La acción humana*, 1949), los autores austriacos de tradición misesiana se basan en la praxeología como herramienta metodológica. La praxeología es la ciencia de la acción humana, cuyas proposiciones se derivan deductivamente del axioma de que el hombre actúa con un propósito. Sus leyes son, por tanto, *a priori* y no requieren contrastación empírica. Ello sitúa a esta tradición en abierta oposición al positivismo y al falsacionismo popperiano, y explica su rechazo del formalismo matemático y de la econometría como instrumentos de conocimiento económico.

A.3. Preferencias sociales (Güth, Schmittberger y Schwarze, 1982; Fehr y Schmidt, 1999)

El *homo economicus* maximiza solo su propio pago, pero los experimentos muestran que muchas personas valoran también los pagos de los demás y la equidad del reparto. En el *juego del ultimátum*, un proponente ofrece un reparto de una suma y el respondedor lo acepta o lo rechaza, en cuyo caso ninguno cobra. La teoría estándar predice la oferta mínima y su aceptación. En los experimentos, en cambio, las ofertas más frecuentes están entre el 40 y el 50 % y las inferiores al 20 % se rechazan aproximadamente la mitad de las veces (Camerer, 2003). Incluso en el *juego del dictador*, sin posibilidad de rechazo, muchos proponentes dan una parte positiva.

Fehr y Schmidt (1999) lo explican con la *aversión a la desigualdad*, según la cual la utilidad de i disminuye con la desigualdad en ambos sentidos, más cuando le es desfavorable,

$$U_i = x_i - \alpha_i \,(x_j - x_i)^+ - \beta_i \,(x_i - x_j)^+, \qquad \beta_i \leq \alpha_i,\ 0 \leq \beta_i < 1$$

donde x_i y x_j son los pagos de i y del otro jugador, $(z)^+$ denota la parte positiva de z (el propio z si es positivo y cero en otro caso), α_i mide la aversión a la desigualdad desfavorable (envidia) y β_i la favorable (culpa). Con un α_i suficientemente alto, rechazar una oferta baja es racional. Las preferencias sociales explican también la cooperación en los juegos de bienes públicos cuando se puede castigar a quien no coopera.

A2. Los economistas clásicos y Marx

Los economistas clásicos y Marx.

Introducción

En el capítulo final de la *Teoría general* (1936), John Maynard Keynes advertía que las ideas de los economistas, tanto cuando aciertan como cuando se equivocan, son más poderosas de lo que suele creerse. De ahí, añadía, que los hombres prácticos que se creen libres de toda influencia intelectual suelan ser esclavos de algún economista difunto. La historia del pensamiento económico estudia precisamente cómo se formaron esas ideas en las sucesivas escuelas, qué preguntas trataron de responder y por qué unas sustituyeron a otras.

La reflexión sobre los fenómenos económicos es muy antigua. Aristóteles ya distinguió el uso y el cambio de los bienes, y siglos después la escolástica medieval y la Escuela de Salamanca discutieron el precio justo, el interés y el valor del dinero. En la Edad Moderna, el **mercantilismo** (1500-1750) identificó la riqueza de las naciones con la acumulación de metales preciosos. Por su parte, la **fisiocracia** (1750-1780) de François Quesnay concibió la economía como un flujo circular sujeto a un orden natural, con la agricultura como única fuente de excedente.

La economía se constituye como ciencia con la **escuela clásica**, a partir de *La riqueza de las naciones* (1776) de Adam Smith y de las obras de Malthus, Ricardo y Mill, y con la crítica de Marx. Esa escuela dejó, sin embargo, problemas sin resolver, en particular una teoría del valor que no explicaba la paradoja del agua y los diamantes y que relegaba la demanda a un papel secundario. Para resolverlos, la **revolución marginalista** de la década de 1870 (Jevons, Menger y Walras) fundó el valor en la utilidad marginal, y de ella nacieron las dos tradiciones neoclásicas dominantes, la de Cambridge, con Marshall y el equilibrio parcial, y la de Lausana, con Walras, Pareto y el equilibrio general.

Ya en el siglo XX, la Primera Guerra Mundial y la Gran Depresión (1929-1933) pusieron en cuestión la capacidad de los mercados para asegurar el pleno empleo. En respuesta, la *Teoría general* de Keynes fundó la **macroeconomía** como cuerpo teórico independiente[1]. La **síntesis neoclásica** (Hicks, Modigliani, Hansen y Samuelson) integró sus ideas en el modelo IS-LM y en la curva de Phillips[2]. Los límites de ese marco fueron, no obstante, haciéndose patentes, pues no lograba explicar el desempleo involuntario como resultado de fallos del mercado, y la estanflación de los años setenta mostró que la curva de Phillips solo se sostenía a corto plazo. Frente a la síntesis, el **monetarismo** de Friedman devolvió al dinero el papel central en las fluctuaciones.

[1] Según la lectura de Leijonhufvud (1968), desarrollada en el tema A04, apartado 3.4.2, lo esencial en Keynes es que la renta (vía multiplicador) se ajusta más deprisa que los precios y el tipo de interés. Ese aspecto se pierde en el análisis de estática comparativa.

[2] La *Teoría general* solo contiene un diagrama (capítulo 14) y no ofrece un análisis sistemático del uso de las políticas monetaria y fiscal. Aun así, el modelo keynesiano del multiplicador se impuso en los años cincuenta porque describía la realidad del momento, se ajustaba a los debates sobre política fiscal y resultaba pedagógicamente útil.

Más tarde, la **nueva macroeconomía clásica** de Lucas incorporó las expectativas racionales y exigió modelos con fundamentos microeconómicos[3], y la **nueva economía keynesiana** aceptó ese método, pero introdujo rigideces nominales, amplificadas por rigideces reales, que devuelven eficacia a la política de estabilización. De la convergencia de ambas corrientes surge la **nueva síntesis neoclásica**, marco de referencia de la macroeconomía actual, cuyos límites volvió a poner a prueba la Gran Recesión de 2008.

La escuela clásica convierte la economía en una disciplina sistematizada, y Marx construye a partir de ella su crítica del capitalismo. Los antecedentes mercantilistas y fisiócratas se recogen en el anexo, y el cuerpo del tema estudia la escuela clásica y a Marx en tres pasos:

- *Formación y madurez*: el pensamiento clásico en las obras de Adam Smith, Thomas Robert Malthus y David Ricardo.

- *Clásicos tardíos*: el pensamiento de Nassau Senior y John Stuart Mill, que prepara la transición al pensamiento neoclásico.

- *Crítica marxista*: las críticas al pensamiento clásico, con especial atención a Karl Marx, a sus implicaciones de política económica y a las razones del fracaso de las economías de planificación central y de la teoría del valor-trabajo.

[3] La macroeconomía keynesiana y de la síntesis partía de relaciones entre agregados en lugar de derivarlas de decisiones individuales. La exigencia de microfundamentos se consolidó a principios de los setenta, tras el acierto de Friedman y Phelps al predecir la aceleración de la inflación. Precedentes: Baumol (1952) y Tobin (1956) sobre el dinero, Ando y Modigliani (1963) sobre el consumo, Jorgenson (1963) sobre la inversión y Muth (1961), que introdujo las expectativas racionales. Hasta Lucas (1972), sin embargo, la agenda se centró en estimar modelos macroeconométricos keynesianos de ecuaciones simultáneas para predicción y simulación de políticas.

1. La formación y maduración del pensamiento clásico: A. Smith, T. R. Malthus, D. Ricardo (s. XVIII a s. XIX)

Entre *La riqueza de las naciones* (1776) y la revolución marginalista de la década de 1870 transcurre casi un siglo de pensamiento económico, que forma el periodo clásico.

Ideas: en general, estos autores comparten la idea de que, gracias a la competencia, el sistema económico se autorregula con mínima intervención estatal, aunque Ricardo subraya el conflicto distributivo entre clases.

1.1. Adam Smith (*La riqueza de las naciones*, 1776)

Formación: catedrático de filosofía moral en la Universidad de Glasgow (1752-1764) y autor de la *Teoría de los sentimientos morales* (1759), conoció a Quesnay y a Turgot en un viaje a Francia (1764-1766).

Importancia: Smith integra en un sistema coherente el conocimiento económico de su generación[4], y suele considerarse que con su obra de 1776 la economía adquiere el estatus de disciplina académica.

Muchas de sus ideas, como la defensa de la libertad económica (su "sistema de libertad natural", próximo al *laissez faire* de los fisiócratas), coinciden con las de los fisiócratas franceses (p. ej. François Quesnay). Smith, sin embargo, ya las defendía antes de su viaje a Francia, y esas ideas deben más a Hutcheson y a David Hume, que advertía del peligro de algunas prácticas mercantilistas[5].

Metodología: en el desarrollo de su obra, Smith combina el razonamiento deductivo con el análisis inductivo de la historia y la observación de la realidad económica de su momento.

1.1.1. Teoría del valor: la paradoja del agua y los diamantes y el valor-coste de producción

Por **valor** se entiende hoy, en general, el precio relativo de un bien[6].

Concepto de valor: en su obra, Smith entendía que el valor tiene dos significados distintos:

- *Valor de cambio* (objetivo): la cantidad de otros bienes que una mercancía permite adquirir, es decir, su precio relativo.

- *Valor de uso* (subjetivo): la aptitud de la mercancía para satisfacer deseos, es decir, su utilidad total.

[4] Smith no fue el primero en escribir un tratado sistemático, pues le precedieron Cantillon (1755) y Steuart (1767), y Schumpeter (1954) llegó a sostener que *La riqueza de las naciones* no contiene ninguna idea, principio o método analítico enteramente nuevo en 1776. Su mérito está en la síntesis y en su influencia.

[5] Hume advirtió que una balanza comercial favorable genera inflación en el país que acumula metales preciosos, lo que reduce sus exportaciones y eleva sus importaciones. En cambio, un aumento gradual de la oferta monetaria estimula la producción mientras dura el ajuste de los precios.

[6] No obstante, esto no ha sido así históricamente. Los primeros economistas-filósofos distinguían entre el concepto de valor y el concepto de precio, entendido el "valor" como la "importancia", "esencialidad" o la "bondad". Por ejemplo, la escolástica, con Santo Tomás de Aquino, se ocupaba del "precio justo", una noción ética del valor ligada a la estimación común.

Smith ilustró esta distinción con dos bienes, el agua y los diamantes. Nada es más útil que el agua, y sin embargo apenas puede obtenerse nada a cambio de ella. Un diamante, en cambio, apenas tiene valor de uso, pero con él pueden adquirirse muchos otros bienes, es decir, su valor de cambio es mucho mayor.

La comparación muestra que el valor de uso de un bien no determina su valor de cambio. Esta divergencia se conoce como *paradoja del agua y los diamantes* o *paradoja del valor*. En las *Lectures on Jurisprudence*, Smith la había explicado por la abundancia o escasez relativa de cada bien, pero en su libro no volvió sobre ella, sino que dejó de lado el valor de uso y centró su teoría en los determinantes del valor de cambio.

La paradoja se debía a que Smith entendía el valor de uso como la utilidad total del bien, no como la utilidad de la última unidad consumida, es decir, su utilidad marginal. Con ello quedaba fuera una dimensión fundamental para la teoría del valor, la escasez relativa de un bien[7]. Hubo que esperar casi cien años a que el análisis marginalista distinguiera entre la utilidad total y la utilidad marginal para resolver la paradoja (tema A03, apartado 1.2.1).

Formación de los precios relativos: Smith la analizó también y distinguió entre:

▪ *Corto plazo*: el precio de mercado es el precio efectivo al que se vende un bien en un momento dado. Depende de la relación entre la cantidad que se lleva al mercado y la *demanda efectiva*, es decir, la de quienes están dispuestos a pagar el precio natural.

▪ *Largo plazo*: con libre competencia (la "perfecta libertad" de Smith, no la competencia perfecta de los neoclásicos), el precio de mercado gravita hacia el *precio natural*, el "precio central" hacia el que tienden de forma permanente los precios de todas las mercancías. Este precio de largo plazo venía determinado por diferentes factores en función del estadio de desarrollo de la sociedad:

- *Sociedad primitiva*: aquella en la que no ha habido acumulación de capital, casi no hay apropiación de tierra y los recursos naturales son bienes libres. En este caso, el valor de cambio o precio relativo de un bien viene dado por la cantidad de trabajo incorporado[8].

- *Sociedad avanzada*: aquella en la que se ha producido acumulación de capital, apropiación de la tierra y los recursos naturales ya no son bienes libres. En este caso, el precio relativo de los bienes viene dado por los costes de producción, es decir, por las remuneraciones derivadas del pago a los factores productivos.

Como medida del valor real, Smith propuso el *trabajo comandado* (*labour commanded*), la cantidad de trabajo ajeno que una mercancía permite comprar, que en la sociedad avanzada supera al trabajo incorporado porque el precio incluye también beneficios y rentas.

No obstante, la teoría del valor basada en el coste de producción en sociedades avanzadas tenía un fallo importante, pues incurría en un razonamiento circular. Los precios relativos se determinan por fuerzas como el salario real, pero, a su vez, el salario real se determina a través de los precios. El análisis neoclásico superó el problema al determinar a la vez, en el equilibrio general, los precios de los bienes y las remuneraciones de los factores, con la productividad marginal como teoría de estas últimas (tema A03).

[7] La explicación del valor por la utilidad y la escasez tenía precedentes en la escolástica tardía y en Galiani (1751). Por ello, Schumpeter (1954) ve en este punto un retroceso de Smith respecto de sus predecesores.

[8] Esto obligó a Smith a plantearse cómo cuantificar el trabajo. Las horas de reloj no bastan como medida, en su opinión, porque dos horas de trabajo no equivalen entre sí si difieren el ingenio y la cualificación que exigen o lo difícil y desagradable de la tarea (es decir, su desutilidad).

1.1.2. Teoría de la distribución de la renta: división de las remuneraciones y doctrina del fondo de salarios

Cómo se reparte la renta total entre los factores es la pregunta que Ricardo situará en el centro de su análisis, pero Smith no le dedica una teoría de la distribución en sentido estricto. Lo que examina es la naturaleza de cada remuneración por separado, y distingue tres categorías.

Salarios: Smith desarrolló a su manera la *doctrina del fondo de salarios*[9], una de las ideas más características del pensamiento clásico. Según esta doctrina, hay un fondo de capital destinado a pagar los salarios. Estas existencias de bienes o capital circulante provienen del ahorro de los capitalistas y se conocen como fondo de salarios[10].

▪ *Corto plazo*: el salario depende de la relación entre los fondos destinados a pagar salarios y la población trabajadora, que la doctrina posterior escribirá como $w = F_w/L$ (donde F_w es el fondo de salarios). Para explicar cómo se fija el salario dentro de ese fondo, Smith combina dos ideas:

- *Teoría de la negociación*: el salario resulta del regateo entre patronos y trabajadores, en el que los primeros, menos numerosos, se coaligan con más facilidad.
- *Crecimiento de la riqueza*: los salarios no dependen del nivel de la riqueza nacional, sino de su ritmo de aumento. Así, son altos donde el capital crece con rapidez (el "estado progresivo") y bajos donde se estanca o declina.

▪ *Largo plazo*: Smith ya fija un mínimo de subsistencia y admite que "la demanda de hombres regula la producción de hombres", y Ricardo convierte ese mínimo en el salario "natural" ($w = w_s$) al añadir la tesis malthusiana de la población.

Beneficios e interés: en la teoría clásica la propiedad y la gestión no están separadas[11]. Por ello, el término beneficio, tal y como lo emplean los clásicos, incluye lo que hoy se denomina beneficios e intereses del capital. Es decir, el interés es solo una parte del beneficio ($i < \pi$).

Smith considera que la competencia en los mercados de bienes, capitales y trabajo hará que, con el tiempo, la tasa de beneficios y, por extensión, el tipo de interés real caigan secularmente (hasta niveles muy bajos en el estado estacionario). Del análisis histórico infiere que existen regiones geográficas en un *estado progresivo* (p. ej. colonias en Norteamérica), con elevados salarios y elevados beneficios, y otras regiones en *estado estacionario* (p. ej. la China de su tiempo). Estas ideas son precursoras de las teorías del crecimiento y desarrollo económico.

El análisis de la naturaleza y fuente de los beneficios fue extraordinariamente breve. Smith aceptó su legitimidad como pago a los capitalistas por desempeñar una función socialmente útil (proveer bienes de subsistencia a los trabajadores y la maquinaria). Sin embargo, también describe la renta y el beneficio como deducciones del producto del trabajo, y de esa ambigüedad partió después la teoría marxista de los beneficios basada en la explotación (apartados 3.2 y 3.3).

Renta: se refiere a la renta de la tierra, que Smith concibe como un precio de monopolio derivado de su apropiación privada. Por ello reprocha a los terratenientes que, "como todos los

[9] La doctrina en sentido estricto, con un fondo dado que se reparte entre los trabajadores ($w = F_w/L$), es posterior y se debe a McCulloch (1825) y a J. S. Mill, que la abandonó en 1869 (apartado 2.2.2). Smith habla solo de los "fondos destinados al pago de salarios", cuyo aumento eleva la demanda de trabajo.

[10] Puesto que producir requiere tiempo, es preciso contar con bienes obtenidos previamente para alimentar, vestir y cobijar a los trabajadores.

[11] Era lógico, porque en aquella época quien aportaba el capital dirigía también el negocio. La sociedad anónima con responsabilidad limitada solo se generaliza a mediados del s. XIX, y la bolsa y los mercados de capitales eran todavía incipientes.

demás hombres, gustan de cosechar donde nunca sembraron", y critica que destinen sus rentas a sostener trabajo improductivo (sirvientes) y gastos suntuarios[12].

Su teoría de la renta, sin embargo, era contradictoria, pues trata la renta unas veces como un componente del precio natural y otras como un efecto del precio. Además, a diferencia de Ricardo, Smith sostiene que el interés de los terratenientes coincide con el interés general, y reserva sus críticas para los comerciantes y fabricantes, que buscan restringir la competencia.

1.1.3. Teoría del crecimiento: determinantes de la riqueza y especialización a través de la división del trabajo

El estudio de Smith sobre la naturaleza y causas de la riqueza de las naciones es una teoría del crecimiento o del desarrollo económico.

Concepto de riqueza, capital y trabajo: el análisis de Smith descansa en tres conceptos:

▪ *Riqueza*: Smith la entiende como un flujo anual de bienes materiales, el "producto anual de la tierra y del trabajo". Esta idea fue precursora de lo que hoy se entiende por producto interior bruto (PIB) o renta. Smith sugirió, además, que la riqueza se midiera en términos per cápita. Se aleja así de los mercantilistas, para quienes la riqueza es la acumulación de metales preciosos que procura el superávit comercial, y de los fisiócratas, para quienes es el excedente agrario.

▪ *Capital*: Smith, al igual que otros autores clásicos, distingue entre capital circulante[13] (que incluye el fondo de salarios) y capital fijo (maquinaria, edificios, mejoras de la tierra).

▪ *Trabajo*: Smith considera que no todo el trabajo es productivo y distingue entre dos tipos:

- *Trabajo productivo*: el que se fija en una mercancía vendible, añade valor al objeto sobre el que se aplica y perdura después de realizado, de modo que puede acumularse, y del que depende la acumulación de riqueza.

- *Trabajo improductivo*: el que no se fija en ningún objeto permanente, como el de los criados, el soberano, el ejército, el clero o las profesiones liberales.

Determinantes de la riqueza: la fábrica de alfileres: según Smith, el producto anual de una nación depende de dos factores, que son los determinantes del crecimiento:

▪ *La productividad del trabajo*: depende de la *división y especialización del trabajo*, idea que Smith recogió de la literatura previa e ilustró con el ejemplo de la fábrica de alfileres[14]. Sobre la división del trabajo, Smith establece dos proposiciones:

- *Fuentes de la productividad*: la división del trabajo eleva la productividad porque aumenta la destreza de cada trabajador, ahorra el tiempo que se perdería al pasar de una tarea a otra y favorece la invención de maquinaria que facilita la labor.

- *Límites*: el grado de división del trabajo depende de dos factores:

[12] "Un hombre se hace rico empleando a una multitud de trabajadores; se hace pobre manteniendo a una multitud de sirvientes" (*La riqueza de las naciones*, 1776).

[13] Smith generaliza esta distinción y le da su terminología moderna, aunque tiene precedentes en los avances anuales y primitivos de Quesnay y Turgot. Que se le llame circulante no quiere decir que sea una variable flujo, sino que se denomina así porque solo rinde beneficio al cambiar de manos o de forma (circulando), mientras que el capital fijo rinde beneficio sin cambiar de dueño.

[14] Smith describe diez trabajadores que, repartiéndose unas dieciocho operaciones, fabricaban más de 48.000 alfileres al día, cuando cada uno por separado no habría hecho ni veinte (1776). El ejemplo procede de la *Encyclopédie* de Diderot y d'Alembert (1755).

- *La extensión del mercado*: solo merece la pena especializarse si hay compradores suficientes para la producción adicional, y por eso un mercado más amplio permite dividir más el trabajo. De ahí procede la importancia que Smith concede al comercio.
- *La acumulación de capital*: la división del trabajo exige una acumulación previa de capital, pues el trabajador solo puede especializarse si dispone de un fondo que lo sostenga y le provea de herramientas y materias primas.

▪ *La proporción entre trabajo productivo e improductivo*: el primero se sostiene con capital y el segundo con renta, de modo que cuanto mayor es el capital respecto de la renta, mayor es la fracción de trabajo productivo y, con ella, el crecimiento.

Ambos determinantes remiten en última instancia a la acumulación de capital, ya que la división del trabajo la presupone y la proporción de trabajo productivo depende de ella[15].

La acumulación de capital procede, a su vez, del ahorro: "la parsimonia, y no la laboriosidad, es la causa inmediata del aumento del capital". En un mundo clásico sin atesoramiento, lo que se ahorra se invierte, es decir, se consume igualmente, pero por trabajadores productivos.

De todo ello se desprenden varias implicaciones:

▪ *Marco institucional*: la acumulación de capital requiere un marco institucional favorable al libre mercado y a la propiedad privada. Con todo, Smith critica la inclinación de algunos comerciantes y fabricantes a procurarse privilegios y monopolios del Estado, pues considera que ello atenta contra el interés general.

▪ *Ahorro de los capitalistas*: el ahorro de los perceptores de beneficios constituye la principal fuente de capital, porque los terratenientes tienden a gastar su renta en lujo y en trabajo improductivo, y los salarios se destinan sobre todo al consumo, aunque sean elevados en el estado progresivo.

▪ *Ahorro y crecimiento*: la idea de que el ahorro es la fuente del crecimiento económico está en la base de los modelos de crecimiento clásicos y, posteriormente, neoclásicos. En el modelo de Solow, sin embargo, un mayor ahorro solo eleva el crecimiento durante la transición al estado estacionario (tema A43, apartado 2.2).

1.1.4. Teoría del comercio: la ventaja absoluta

Smith entra en la teoría del comercio sobre todo para rebatir a los mercantilistas (Anexo A.1). Estos veían el comercio exterior como un juego de suma cero y recomendaban atesorar oro y evitar los déficits comerciales mediante el proteccionismo. Para Smith, en cambio, el libre comercio, tanto exterior como interior, es de suma positiva, y su rechazo de las barreras comerciales descansa en el concepto de ventaja absoluta.

Idea: un país goza de *ventaja absoluta* en un bien cuando lo produce con un coste unitario inferior al de cualquier otro país o, lo que es lo mismo, con un producto medio del trabajo (Pme_L) superior.

Razonamiento: siempre que la relación real de intercambio quede entre los precios relativos de autarquía de ambos, cada país gana especializándose por completo en el bien en que tiene esa ventaja y exportándolo. De ahí la máxima de que "todo padre de familia prudente" nunca

[15] "El número de trabajadores productivos nunca puede incrementarse mucho si no es como consecuencia de un aumento del capital o de los fondos destinados a mantenerlos. […] En cualquiera de los dos casos, es casi siempre necesario más capital". Esto además resta importancia al papel de los terratenientes.

produce en casa lo que le cuesta más hacer que comprar. La causa del comercio es, por tanto, la posibilidad de ganancias mutuas, y el patrón de comercio se explica por las diferencias tecnológicas absolutas.

La apertura, además, da salida a la producción que el mercado interior no absorbe (*vent for surplus*, la expresión de J. S. Mill que recuperó Myint, 1958) y amplía el mercado, de modo que aumenta la división del trabajo, eleva el crecimiento y retrasa el estado estacionario.

Conclusiones:

- *Aspectos positivos*: es una primera aproximación a la explicación de las causas del comercio internacional y el patrón de comercio.

- *Limitaciones*: no explica por qué países con ventaja absoluta en todos los sectores participan en el comercio internacional, ni cómo se fija la relación real de intercambio, y solo admite la especialización total por considerar costes unitarios constantes (una frontera de posibilidades de producción lineal).

> La teoría de la ventaja absoluta, con su ejemplo numérico, se desarrolla en el tema B05, apartado 1.1, como antecedente de la teoría neoclásica del comercio internacional.

1.1.5. Teoría monetaria y bancaria: la doctrina de las letras o billetes reales y el mecanismo flujo-especie

Smith acepta la teoría cuantitativa en una economía cerrada o con papel moneda inconvertible. En una economía pequeña y abierta con moneda convertible, en cambio, la cantidad de dinero es endógena, pues se ajusta al valor de las mercancías que han de circular y el exceso sale al exterior[16]. El papel moneda bancario, al sustituir al costoso dinero metálico (la "gran rueda de la circulación"), libera además capital para usos productivos y fomenta la división del trabajo.

Para evitar la sobreemisión, Smith propone un sistema bancario basado en la **doctrina de los billetes reales**, según la cual los bancos deben emitir billetes únicamente contra letras comerciales reales (respaldadas por transacciones comerciales efectivas). Se busca así que la emisión se ajuste a las necesidades del comercio y no exceda el valor del oro y la plata que circularían en ausencia de papel moneda.

Mecanismo flujo-especie: Smith no menciona en su libro el mecanismo que había formulado su amigo David Hume (1752), pero sí en los apuntes de alumnos de sus clases en Glasgow a principios de la década de 1760, publicados póstumamente.

El mecanismo establece que, si se cumple la teoría cuantitativa, los intercambios son solo comerciales y rige el patrón oro, cualquier desequilibrio de la balanza de pagos se corrige automáticamente a través de los precios. En concreto, el país con déficit pierde oro, su masa monetaria y sus precios caen y su depreciación real abarata sus exportaciones, mientras que el país con superávit experimenta lo contrario.

> El mecanismo flujo-especie de Hume, con sus etapas y sus limitaciones, se desarrolla en el tema B12, apartado 1.1.1.

[16] Véanse Humphrey (1981) y Laidler (1981). Según estos autores, Smith no incorpora a *La riqueza de las naciones* el mecanismo flujo-especie de Hume, porque en una economía pequeña y abierta el arbitraje de mercancías liga los precios interiores a los mundiales.

1.1.6. Implicaciones de política económica: la mano invisible y el *laissez faire*

▪ *Laissez faire*: la principal conclusión sobre política económica es que el Gobierno debía seguir una política de *laissez faire*[17]. Así, le corresponden tres funciones: la defensa, la administración de la justicia y las obras e instituciones públicas que la iniciativa privada no provee (incluida la educación).

▪ *Mano invisible*: la tesis del *laissez faire* descansa en la idea de que, cuando el mercado funciona sin regulación, la búsqueda del interés privado desemboca en el interés público. Para Smith, nadie juzga mejor que cada individuo su propio interés (es el mejor juez de él), de modo que debe ser libre para perseguirlo. Con esa libertad, cada persona no solo logrará su propio interés, sino que, guiada por una mano invisible y sin que nadie coordine las decisiones individuales, promoverá también el bienestar colectivo[18].

▪ *Hacienda pública*: Smith quiere costear esas tres funciones con impuestos sujetos a cuatro *cánones*, la igualdad (en proporción a la renta de cada uno), la certeza, la comodidad en el pago y la economía en la recaudación. Su tributo preferido es el que grava la renta de la tierra, cuyo efecto sobre la actividad productiva no es directo. Rechaza, en cambio, el endeudamiento, porque, al pedir prestado a cambio de un interés, el Estado absorbe fondos que de otro modo se habrían invertido en producir, y con ello convierte trabajo productivo en improductivo.

1.1.7. Conclusiones

Aspectos positivos: sobre las ideas de la mano invisible y el *laissez faire* se asienta el liberalismo económico, que ha influido en la política económica mundial[19].

Limitaciones: la mano invisible es una intuición y no una demostración. Su formalización posterior, el primer teorema del bienestar (tema A22, apartado 2.1), muestra que el resultado exige competencia perfecta y ausencia de fallos de mercado, y que garantiza la eficiencia paretiana pero no la equidad. En la teoría del valor, además, quedaron sin resolver la paradoja del agua y los diamantes y la circularidad del valor-coste.

1.2. Thomas Robert Malthus (*Ensayo sobre la población*, 1798)

Formación: clérigo anglicano formado en matemáticas en Cambridge, ocupó desde 1805 la primera cátedra de economía política de Inglaterra, en el colegio de la Compañía de las Indias Orientales.

Importancia: hasta la obra de Ricardo (1817), el pensamiento económico estuvo dominado por Smith. El *Ensayo* de Malthus influyó, sin embargo, en Ricardo, con quien su autor mantuvo una larga polémica sobre la renta, el valor y el exceso general de oferta.

[17] Smith no emplea la expresión, que procede de los fisiócratas, sino la de "sistema de libertad natural", y admite excepciones expresas (Viner, 1927). Entre ellas figuran las leyes de navegación, porque "la defensa es mucho más importante que la opulencia", un tope legal al tipo de interés, la prohibición de los billetes bancarios de pequeño importe y la educación elemental pública.

[18] En *La riqueza de las naciones*, la expresión aparece una sola vez (libro IV, cap. 2), a propósito del inversor que prefiere emplear su capital en la industria nacional antes que en la extranjera, y antes en la *Teoría de los sentimientos morales* (1759). Su lectura como principio general de coordinación por el mercado es posterior. Además, se ha hecho una utilización abusiva de la metáfora de la mano invisible, puesto que Smith en 1776 ya explicó con claridad el peso de los factores institucionales como condicionantes de la actuación de los individuos.

[19] Smith no defendía un capitalismo salvaje. El individuo que persigue su propio interés es el mismo sujeto moral que retrata en su obra de 1759.

Metodología: Malthus pone más énfasis que Ricardo en el razonamiento inductivo, porque la evolución observada de las fuerzas demográficas es fundamental en su análisis.

1.2.1. Teoría del crecimiento: la tesis sobre la población

Las aportaciones de Malthus sobre la población se recogen en el *Ensayo sobre el principio de la población* (1798; 2.ª ed., ampliada, 1803). Tres circunstancias explican la tesis malthusiana:

- *Crecimiento de la población*: en Inglaterra se constataba que el número de habitantes presionaba sobre los alimentos disponibles.

- *Pobreza creciente*: afectaba a los grupos de menor renta y encarecía la asistencia de las leyes de pobres.

- *Ideología*: Godwin y Condorcet atribuían la miseria a las instituciones y creían en una sociedad perfectible sin límite. Malthus quiere refutarlos y sostiene que la pobreza obedece a una ley natural, la propensión de la población a aumentar más rápido que los medios de subsistencia.

Tesis sobre la población: según Malthus, una población sin frenos se duplica cada veinticinco años y sigue, por tanto, una progresión geométrica, mientras que los alimentos aumentan a lo sumo en progresión aritmética. La presión del número sobre los alimentos disponibles condena así a los estratos de renta baja a la pobreza y el sufrimiento.

Como Malthus no contempla que la producción de alimentos mejore al ritmo de la población, considera que existen frenos que ajustan el crecimiento de la población al de los alimentos. Concretamente, distingue entre:

- *Frenos positivos*: hambres, guerras, epidemias y, en general, todo lo que eleva la mortalidad.

- *Frenos preventivos*: actúan sobre la tasa de natalidad y la reducen. El que Malthus defendía, desde la segunda edición (1803), era la restricción moral (*moral restraint*), es decir, el retraso del matrimonio y la abstinencia sexual previa a este. Por el contrario, clasificaba los métodos anticonceptivos dentro del "vicio" y los rechazaba, mientras que los neomalthusianos posteriores los reivindicaron.

1.2.2. Teoría de los ciclos: la insuficiencia de demanda y el subconsumismo malthusiano

En sus *Principios de economía política* (1820), Malthus estudió también las consecuencias del ahorro y la acumulación de capital, y se preguntó qué ocurre si la demanda agregada de bienes finales es menor que la oferta agregada. La pregunta desafiaba a los demás clásicos, que daban por supuesta la ley de Say (1803), según la cual no puede haber un exceso general de oferta, aunque sí excesos parciales (tema A04, apartado 1.3.3). La motivación de Malthus fue la crisis de sobreproducción[20] que siguió en Inglaterra a las Guerras Napoleónicas (1803-1815).

Idea: Malthus distingue entre los bienes básicos, cuya producción genera su propia demanda, y el resto de bienes, para los que sí puede existir sobreproducción. La causa es que la demanda de artículos de lujo y comodidad depende de los hábitos de gasto de los grupos improductivos de la sociedad, como los terratenientes.

Razonamiento: sigue tres pasos:

[20] La economía británica había ampliado su capacidad productiva para atender las necesidades de la guerra.

1. *Demanda de los trabajadores*: el trabajo no percibe todo el producto, de modo que la demanda de los trabajadores no basta para adquirir todos los bienes finales.

2. *Ahorro de los capitalistas*: los capitalistas tienen poder adquisitivo suficiente, pero no voluntad de consumir, porque ahorran. Si ese ahorro se destina a acumular capital, amplía la capacidad productiva sin una demanda efectiva que absorba la mayor oferta.

3. *Consumo improductivo*: dado que la demanda efectiva de trabajadores y capitalistas es insuficiente, la diferencia debe cubrirla el consumo de quienes consumen y no producen, como los funcionarios públicos, los maestros, los sirvientes y los terratenientes. En caso contrario, la economía no alcanza el pleno empleo, lo que implica rechazar la ley de Say.

Conclusiones:

- *Aspectos positivos*: Malthus fue uno de los primeros clásicos que plantearon la posibilidad de un exceso general de oferta por insuficiencia de la demanda efectiva, idea que Keynes (1936) recuperó como antecedente de la suya.

- *Limitaciones*: no explicó con claridad por qué la demanda efectiva había de ser insuficiente (Keynes, 1936). Además, como admitía que todo lo ahorrado se invierte, su argumento es de exceso de capacidad y no de atesoramiento, lo que lo separa de Keynes, y su defensa del consumo improductivo de los terratenientes servía para justificar sus rentas.

1.2.3. Implicaciones de política económica: la ley de hierro de los salarios

- *Leyes de pobres*: en la tesis malthusiana se apoyó la *ley de hierro de los salarios*, nombre que acuñaría después Ferdinand Lassalle (1863). Su lógica es que, como cada subida del salario real hace crecer la población, cualquier esfuerzo por mejorar la situación económica de los grupos con rentas bajas está condenado de antemano al fracaso. Por ello, Malthus criticó abiertamente las leyes de pobres vigentes en Inglaterra desde 1601 y ampliadas con el sistema de Speenhamland (1795), y propuso su abolición gradual, porque a su juicio fomentaban la natalidad sin aumentar los alimentos.

- *Política comercial*: a diferencia de Ricardo, Malthus (1815) defendió las leyes de los cereales. Sus razones eran la seguridad del abastecimiento y el papel de los terratenientes como consumidores.

1.2.4. Conclusiones

Aspectos positivos: la tesis de la población suscitó polémica y alcanzó una repercusión muy superior a la que su autor había previsto, y Ricardo la incorporó a la economía política clásica.

Limitaciones: sus presagios no se cumplieron en los países industrializados, porque el progreso técnico en la agricultura elevó la producción de alimentos más deprisa que la población y la transición demográfica redujo la natalidad al aumentar la renta. Esos presagios (el "fantasma malthusiano") contribuyeron, con todo, a que la expresión "ciencia lúgubre" (*dismal science*), acuñada por Thomas Carlyle en 1849 en otro contexto, quedase asociada popularmente a la economía.

1.3. David Ricardo (*Principios de economía política y tributación*, 1817)

Formación: agente de bolsa sin estudios universitarios, hizo fortuna con la deuda pública de las Guerras Napoleónicas y fue diputado desde 1819. Se interesó por la economía tras leer *La riqueza de las naciones* en 1799.

Importancia: Ricardo hace de la economía un saber abstracto y de alcance general. En su obra principal, *Principios de economía política y tributación* (1817), construye el primer sistema analítico de la economía política, hasta tal punto que se le tiene por precursor de la elaboración de modelos económicos.

Ricardo escribe en la época de las Guerras Napoleónicas (1803-1815). Al interrumpir el comercio, el conflicto obligó a Gran Bretaña a producir más alimentos. Cuando los precios agrícolas bajaron desde su máximo de 1812, los terratenientes obtuvieron las leyes de los cereales[21], que prolongaban en la paz el amparo de los años de guerra. Muchos economistas las combatieron, pero las leyes siguieron en vigor hasta 1846. Ricardo construye su teoría de la distribución para fundamentar la crítica a esas leyes, primero en su ensayo de 1815 y después en su libro de 1817. En esta segunda obra, la teoría de la distribución se apoya en una reformulación de la teoría del valor de Smith.

Metodología: Ricardo razona de forma estrictamente deductiva, a partir de unos pocos supuestos simplificadores (rendimientos decrecientes en la agricultura, una población malthusiana y, en la versión simplificada habitual de su modelo, un único bien salario).

1.3.1. Teoría del valor: la teoría del valor-trabajo al 93 %

Concepto de valor: para entender la aportación de Ricardo en cuanto al valor hay que tener en cuenta tres cuestiones:

- *Valor de uso y valor de cambio*: acepta la distinción de Smith y precisa que la utilidad no es la medida del valor de cambio, aunque es imprescindible para que exista.

- *Variaciones del valor*: le interesan sobre todo las variaciones de los precios relativos en el tiempo, y busca una *medida invariable del valor* que las separe de los cambios en la distribución[22].

- *Distinción entre bienes*: distingue entre bienes libremente reproducibles, cuyo valor depende de la cantidad de trabajo necesaria para obtenerlos, y bienes no reproducibles (p. ej. estatuas y cuadros raros o vinos de calidad singular), cuyo valor solo deriva de la escasez. Dado que estos últimos representan una proporción pequeña de la economía, se centra en los primeros.

Formación de los precios relativos: a diferencia de Smith, Ricardo deja fuera la renta de la tierra cuando explica de qué dependen los precios en el largo plazo, y en ello consiste lo más

[21] En inglés, Corn Laws. La ley de 1815 prohibía importar trigo mientras el precio interior no alcanzara 80 chelines por *quarter*. Desde 1828, la prohibición se sustituyó por aranceles móviles.

[22] Al medir el valor por el trabajo incorporado, y no por la suma de salarios, beneficios y rentas, evita el razonamiento circular de Smith (apartado 1.1.1). Ricardo no encontró esa medida, porque ninguna mercancía real se produce siempre con la misma proporción de trabajo y capital. Sraffa (1960) construyó con ese fin una mercancía compuesta, la *mercancía patrón*, cuyo valor no depende de cómo se reparta el producto entre salarios y beneficios.

novedoso de su teoría del valor. En concreto, para Ricardo la renta de la tierra no forma parte del coste de producción[23].

Por tanto, según Ricardo, a largo plazo el valor de cambio (precio) de cualquier bien que pueda reproducirse libremente y en competencia depende del coste de producción asociado al empleo de trabajo y de capital. Como el capital es, en definitiva, reducible a trabajo acumulado, el trabajo incorporado se convierte en el principal determinante del precio natural[24].

Ahora bien, Ricardo admitía otros determinantes, como la importancia relativa del capital fijo, una segunda causa de valor[25], o las variaciones de los salarios, que alteran los precios relativos cuando las proporciones de capital difieren entre sectores. Con todo, pensaba que sus efectos eran pequeños y que la cantidad de trabajo era el determinante principal de las variaciones de los precios relativos. Por eso George Stigler (1958a) llamó a la teoría de Ricardo "teoría del valor-trabajo al 93 %".

1.3.2. Teoría de la distribución de la renta: la teoría de la renta diferencial

La teoría de la distribución es una de las principales aportaciones de Ricardo. Se apoya en la división de las remuneraciones de Smith y en el concepto de renta diferencial.

Salarios: a corto plazo, Ricardo explica el salario real con la doctrina del fondo de salarios. A largo plazo entra en juego la tesis malthusiana, porque los movimientos de la población tienden a devolver el salario de mercado a su nivel natural, aunque, en una sociedad que progresa, puede mantenerse por encima de él "durante un periodo indefinido"[26]. La versión simplificada habitual del modelo supone que el trabajador gasta casi todo su salario en trigo, pero Ricardo vincula el salario natural al precio de "los alimentos, bienes necesarios y comodidades" que el trabajador requiere.

Beneficios e interés: el beneficio se obtiene por diferencia, al descontar del producto la renta y los salarios. Su caída secular, que Smith achacaba a la competencia, se debe para Ricardo a los rendimientos decrecientes agrícolas (punto 4). Según la lectura de Sraffa (1951), no compartida por todos los historiadores, en el ensayo de 1815 la tasa de beneficios agrícola gobierna la del conjunto de la economía.

Renta: el núcleo de su teoría de la distribución es la definición de la renta como "la parte del producto de la tierra que se paga al terrateniente por el uso de las facultades originarias e indestructibles del suelo"[27]. La renta diferencial o renta ricardiana es aquella que iguala la tasa de beneficio de tierras con diferente fertilidad. Con iguales dosis de capital y trabajo, es la diferencia entre el producto de la tierra fértil, A, y el de la marginal, B, que no paga renta, $R_A = Y_A - Y_B$, de modo que, pagada la renta, el beneficio es el mismo en ambas.

[23] Para Ricardo, el precio de un quintal de trigo es su coste marginal, es decir, el del quintal producido en las peores condiciones. El precio se forma en ese margen, donde no se paga renta. Es, por tanto, el precio el que explica la renta, y no a la inversa. En sus palabras, "el trigo no es caro porque se pague una renta, sino que se paga una renta porque el trigo es caro". Por eso el hecho de que tierras de fertilidad distinta paguen rentas distintas no altera cómo evolucionan con el tiempo los precios relativos.

[24] Smith ceñía la teoría del trabajo incorporado a la sociedad primitiva. Ricardo la extiende a cualquier estadio de la economía.

[25] El capital fijo tarda más en recuperarse y esa espera debe remunerarse. Por ello, a igualdad de trabajo incorporado, el precio relativo de un bien aumenta con la ratio capital fijo / capital circulante empleada en producirlo.

[26] Para Ricardo, el salario de subsistencia no es un mínimo fisiológico fijo, sino que varía según la época y las costumbres.

[27] La teoría de la renta diferencial se publicó casi a la vez en 1815, en los folletos de Malthus, West, Torrens y el propio Ricardo, y tiene un antecedente en Anderson (1777).

De este concepto deduce Ricardo que la renta retribuye la simple propiedad de la tierra, no un servicio útil a la sociedad. Este argumento lo utiliza para defender la abolición de las leyes de los cereales:

1. *Calidad de las tierras*: varía de las tierras fértiles a las estériles (por simplicidad, se consideran solo los dos extremos). Todas presentan rendimientos (marginales) decrecientes.

2. *Márgenes intensivo y extensivo*: el intensivo refleja el producto marginal de sucesivas dosis de factores sobre una misma parcela, y el extensivo, el de añadir nuevas parcelas (p. ej. del valle a la ladera).

3. *Leyes de los cereales*: al restringir la importación, fomentan la explotación local, lo que agota los márgenes intensivos y empuja a cultivar tierras de menor fertilidad (margen extensivo).

4. *Rendimientos decrecientes*: conforme se intensifica el cultivo en las buenas tierras y se extiende a las peores, los rendimientos decrecientes elevan el coste y el precio del grano. Esto tiene dos efectos:

- *Renta diferencial*: aumenta en las tierras más fértiles (se amplía la diferencia respecto de la tierra marginal).

- *Salario de subsistencia*: sube en términos monetarios, porque, para mantener su poder adquisitivo en trigo, los trabajadores necesitan un salario monetario mayor. Con el salario real constante en el nivel de subsistencia, la mayor participación de los salarios en el valor del producto comprime el beneficio.

Resultados:

- *Renta ricardiana*: la mera propiedad de un recurso limitado en cantidad y desigual en calidad genera una renta diferencial en favor de los terratenientes.

- *Leyes de los cereales*: al extender el cultivo a tierras menos fértiles, elevan las rentas de los terratenientes y aceleran la llegada al estado estacionario.

- *Caída del beneficio*: el aumento de la población eleva las rentas y reduce progresivamente el beneficio, que tiende hacia un mínimo, y de la evolución de las remuneraciones en el tiempo Ricardo deriva una "teoría del crecimiento".

Gráficamente, la teoría ricardiana de la renta se representa en el plano trabajo-salario (figura A2.1):

1. *Productos medio y marginal del trabajo*, $\mathrm{Pme}_L(L)$ y $\mathrm{Pmg}_L(L)$: decrecen por los rendimientos decrecientes, y el marginal queda por debajo del medio, porque cada dosis adicional rinde menos que las anteriores.

2. *Salarios*: el de subsistencia, w_s, es horizontal, fijado por la tesis malthusiana. El de mercado lo supera mientras dura la acumulación, y su senda, $w(t)$, converge a w_s al crecer la población y caer el beneficio.

3. *Reparto antes del estado estacionario* (empleo L_0 y salario w_0): el producto, $\mathrm{Pme}_L(L_0)\,L_0$, se reparte en tres rectángulos. La renta, $R_0 = [\mathrm{Pme}_L(L_0) - \mathrm{Pmg}_L(L_0)]\,L_0$, va del producto medio al marginal; el beneficio, $\pi_0 = [\mathrm{Pmg}_L(L_0) - w_0]\,L_0$, del marginal al salario, y la masa salarial, $F_{w,0} = w_0\,L_0$, queda bajo el salario.

4. *Estado estacionario*: el empleo L^* se fija donde el producto marginal iguala al salario de subsistencia, y el beneficio se anula[28].

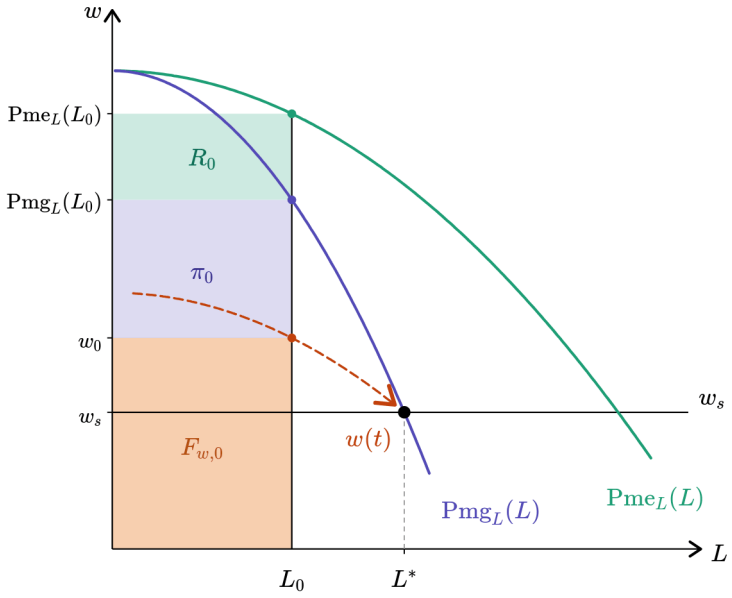

Figura A2.1: Teoría ricardiana de la renta y la distribución. Antes del estado estacionario, con empleo L_0 y salario de mercado w_0, el producto se reparte entre la renta, R_0, el beneficio, π_0, y los salarios, $F_{w,0}$. El salario de mercado, $w(t)$, converge al de subsistencia, w_s, y en el estado estacionario, L^*, donde $\text{Pmg}_L(L^*) = w_s$, el beneficio se anula.

1.3.3. Teoría del crecimiento: el estado estacionario "lúgubre"

Ricardo comparte con Smith la idea de que el crecimiento lo impulsa la acumulación de capital, alimentada por los beneficios. Cuando estos se agotan, la economía desemboca en un **estado estacionario**. En él el beneficio cae hasta un mínimo que ya no compensa la molestia y el riesgo de invertir, de modo que la acumulación se detiene, el salario sigue anclado en la subsistencia y los terratenientes se quedan, en forma de renta, con todo lo demás.

En la tercera edición de los *Principios* (1821), el capítulo "Sobre la maquinaria" admite que la sustitución de trabajo por máquinas puede perjudicar a los trabajadores, idea que Marx retomará (apartado 3.4).

1.3.4. Teoría del comercio: la ventaja comparativa

David Ricardo amplía la visión de Smith al introducir, con el ejemplo del paño y el vino de Inglaterra y Portugal, el concepto de **ventaja comparativa**, que descansa en el concepto de coste de oportunidad[29].

[28] El modelo simplificado iguala a cero el beneficio mínimo. En Ricardo, la acumulación se detiene antes, cuando el beneficio es tan bajo que ya no compensa la molestia y el riesgo de invertir (apartado 1.3.3).

[29] Esta es la interpretación moderna, debida a Haberler (1930), que reformuló la ventaja comparativa en términos de coste de oportunidad con ideas de la escuela austriaca. El concepto de coste de oportunidad apareció después

Idea: un país tiene ventaja comparativa en la producción de un bien si el coste de oportunidad de producirlo, en términos del otro bien, es menor que en otros países.

Razonamiento: aunque un país tenga ventaja absoluta en la producción de ambos bienes, existen ganancias mutuas del comercio siempre que la relación real de intercambio se sitúe entre los precios relativos de autarquía. En el ejemplo de Ricardo, ese país es Portugal, que producía el paño y el vino con menos trabajo que Inglaterra.

La causa del comercio sigue siendo la posibilidad de ganancias mutuas, pero el patrón de comercio se explica por las diferencias tecnológicas relativas, y cada país se especializa totalmente en el bien en el que tiene ventaja comparativa. Si ambos países tienen los mismos costes de oportunidad, el comercio no beneficia a ninguno.

Conclusiones:

- *Aspectos positivos*: Ricardo resolvía el problema de la desventaja absoluta, porque una industria con ventaja comparativa puede competir aunque su país esté más atrasado tecnológicamente.

- *Limitaciones*: no determina la relación real de intercambio, mantiene la especialización total y descansa en un único factor y en la teoría del valor-trabajo, de modo que no explica el origen de las diferencias de productividad.

> La teoría de la ventaja comparativa, con su ejemplo numérico y el cálculo de los costes de oportunidad, se desarrolla en el tema B05, apartado 1.2 (y su formalización, en el anexo A.2), como antecedente de la teoría neoclásica del comercio internacional.

1.3.5. Teoría monetaria y bancaria: el debate bullonista y las virtudes del papel moneda

Por lo que respecta a su teoría monetaria, Ricardo admite las ventajas del papel moneda, que en su época circulaba sin convertibilidad en oro (periodo de restricción del Banco de Inglaterra, 1797-1821), aunque advierte a la vez de los riesgos que ello entraña.

Durante la restricción, el precio del oro subió por encima de su paridad y la libra se depreció. En el *debate bullonista*, Ricardo (1810) atribuyó ambos hechos al exceso de emisión del Banco de Inglaterra, frente a los antibullonistas, que invocaban la doctrina de los billetes reales. Apoya así la teoría cuantitativa, en la tradición de Locke y Hume, y sus tesis, junto con las de Thornton, inspiraron el *Bullion Report* (1810)[30].

Ricardo propone, como remedio frente al riesgo de emisión, ligar el *stock* de oferta monetaria a un respaldo material, por ejemplo los lingotes de oro. La propuesta permite ver en Ricardo a un precursor del patrón oro en su modalidad de patrón lingote-oro (el llamado *ingot plan*, esbozado en 1811 y desarrollado en 1816). Su plan de 1824, que separaba la emisión de la banca, inspiró a la escuela monetaria (apartado 2.2.5).

de Ricardo en la literatura (el término lo acuñó D. I. Green en 1894). Robert Torrens (1815) había apuntado la idea poco antes, aunque la formulación completa es de Ricardo (Ruffin, 2002).

[30] Lo redactaron Francis Horner, William Huskisson y Henry Thornton, miembros del comité de la Cámara de los Comunes. Ricardo, que aún no era diputado, influyó con sus cartas y su folleto de 1809-1810.

1.3.6. Implicaciones de política económica: el *laissez faire*

La esencia de su política económica, al igual que en Smith, se encuentra en el *laissez faire*, que Ricardo concreta en dos ámbitos:

- *Estado estacionario*: para retrasar la llegada al estado estacionario, Ricardo propone el libre comercio de cereales, que permite importar alimentos baratos, y confía en las mejoras técnicas en la agricultura. El libre comercio y las mejoras técnicas frenan el alza del precio del grano y del salario monetario, sostienen el beneficio y fomentan la acumulación de capital.

- *Hacienda pública*: como Smith, Ricardo prefiere financiar el gasto público con impuestos. En su ensayo de 1820 concluye que, en teoría, financiar un gasto con impuestos o con deuda es equivalente, pero duda de que se cumpla en la práctica, porque los contribuyentes no perciben la carga futura de la deuda. Barro (1974) revivió la idea, y Buchanan (1976) la bautizó como *equivalencia ricardiana* (tema A06, apartado 3.2.1). En la parte de tributación de los *Principios*, Ricardo analiza además quién soporta cada impuesto. Un impuesto sobre la renta de la tierra recae sobre el terrateniente y no altera el precio del grano, mientras que uno sobre los salarios eleva el salario monetario y recae sobre el beneficio.

1.3.7. Conclusiones

Aspectos positivos: la teoría de Ricardo es el primer modelo coherente que explica cómo varía con el tiempo la parte del ingreso que corresponde a trabajadores, capitalistas y terratenientes (Smith se había limitado a describir cada remuneración). La teoría ricardiana pone de manifiesto, además, que los intereses de las distintas clases sociales se contraponen, algo que influirá en cierta medida en el pensamiento posterior de Marx.

Limitaciones: su método abstracto, que Schumpeter (1954) llamó *vicio ricardiano*, extrae conclusiones de política de supuestos muy simplificados. Además, la predicción del estado estacionario no se cumplió, pues descansaba en la tesis malthusiana de la población y en unos rendimientos decrecientes que el progreso técnico en la agricultura compensó con creces.

1.4. Conclusiones

Smith, Malthus y Ricardo explican el crecimiento del producto y su reparto entre las clases en una economía que, gracias a la competencia, se autorregula con mínima intervención estatal. Smith sitúa el origen del crecimiento en la división del trabajo, limitada por el mercado y por el capital acumulado, y confía en que el interés privado, en libre competencia, promueva el bienestar colectivo. Malthus opone a ese optimismo la presión de la población sobre los alimentos y el riesgo de que el ahorro deje producción sin demanda efectiva.

Ricardo integra la población malthusiana y los rendimientos decrecientes en el primer sistema analítico de la economía política. La extensión del cultivo a tierras peores eleva en él la renta y comprime el beneficio hasta detener la acumulación, desenlace que el libre comercio de cereales y el progreso técnico solo retrasan. Su ventaja comparativa extiende, además, la ganancia mutua del comercio a los países sin ventaja absoluta.

2. La transición al pensamiento neoclásico: el pensamiento clásico tardío (s. XIX)

Inmediatamente después de Ricardo, tanto la teoría económica como el propio sistema capitalista recibieron las críticas de humanistas y socialistas, entre ellos los llamados socialistas ricardianos, que usaron la teoría del valor-trabajo para reclamar para el trabajo todo el producto. En este contexto surgen las aportaciones de Nassau Senior y John Stuart Mill, que revisan el sistema ricardiano y preparan la transición al pensamiento neoclásico.

2.1. Nassau Senior (*Esbozo de la ciencia de la economía política*, 1836)

Formación: abogado y primer titular de la cátedra Drummond de economía política de la Universidad de Oxford (1825-1830 y 1847-1852), Senior redactó con Edwin Chadwick el informe de 1834 que dio lugar a la nueva ley de pobres.

Metodología: entre las aportaciones más destacadas de Nassau Senior, recogidas en su obra de 1836, se encuentran las relativas al ámbito y la metodología de la economía. Senior figura entre los primeros economistas que defendieron sin ambigüedad que la economía ha de ser una ciencia positiva, separada del consejo político, porque sus conclusiones no autorizan al economista a añadir "una sola sílaba de consejo". La distinción entre economía positiva, economía normativa y arte la sistematizará después J. N. Keynes (1891). Deduce la ciencia de cuatro proposiciones elementales: el deseo de obtener la mayor riqueza con el menor sacrificio, el principio de la población, la mayor productividad del trabajo cuando sus productos se emplean como medios de producción (el capital) y los rendimientos decrecientes agrícolas.

Teoría del valor y del interés: Senior hace depender el valor de la utilidad y de la limitación de la oferta (y enuncia la utilidad decreciente), y el coste de producción, que para él incluye el trabajo y la abstinencia, actúa a través de esa limitación. Su teoría del interés basada en la abstinencia, la remuneración por renunciar al consumo inmediato, es precursora de la espera (*waiting*) de Marshall y de las teorías del interés basadas en la preferencia temporal[31]. Con esa teoría, Senior trata de superar las deficiencias de la teoría clásica sobre la tasa de beneficios, que no distingue una teoría de los beneficios de una teoría del interés.

2.2. John Stuart Mill (*Principios de economía política*, 1848)

Formación: educado por su padre, James Mill, amigo de Bentham y de Ricardo, John Stuart Mill recibió desde la infancia una formación excepcionalmente vasta (griego, lógica, literatura clásica, química…).

Importancia: fue un pensador singular y ecléctico. Se le suele ver como el puente entre clásicos y neoclásicos, ya que desarrolló varios aspectos de la teoría heredada y adelantó intuiciones sobre el bienestar y sobre el ciclo económico que otros desarrollarían después. Hasta finales del s. XIX, sus *Principios de economía política* (1848) fueron el texto de referencia de la disciplina.

Metodología: Mill razona de forma deductiva. En su ensayo de 1836 concibe la economía como una ciencia abstracta que razona *a priori* a partir de un agente que solo desea la riqueza (el futuro *homo economicus*).

[31] En sus *Cartas sobre la Ley de Fábricas* (1837), Senior sostuvo que todo el beneficio neto de una fábrica textil se obtenía en la última hora de la jornada, de modo que acortarla arruinaría a la industria. Marx ridiculizó en *El capital* ese argumento de la "última hora", como también la abstinencia como justificación del beneficio.

2.2.1. Teoría del valor: revitalización del valor-coste de producción

Concepto de valor: Mill se presenta como continuador de Ricardo. Para él, el valor de los bienes reproducibles depende del coste de producción, formado por salarios y beneficios, del que la cantidad de trabajo es el elemento principal. Con todo, reconoce que la demanda influye en los precios, aunque nunca representó la oferta y la demanda mediante curvas.

Formación de los precios relativos: Mill distinguió tres tipos diferentes de bienes[32]:

1. *Bienes de oferta inelástica*: el precio depende solo de la demanda, como ocurre con las obras de arte[33].

2. *Bienes de oferta totalmente elástica*: en las manufacturas, por ejemplo, la oferta fija el precio, que refleja los costes de producción. Mill supone además que esos costes son constantes, de forma que producir más no altera el coste marginal.

3. *Bienes de oferta con pendiente positiva*: oferta y demanda determinan juntas el precio, como en la agricultura, porque los costes crecen al regir la ley de rendimientos marginales decrecientes.

Los marginalistas se apoyaron más tarde en esta clasificación para objetar que la oferta determina por completo el precio solo en los bienes del segundo tipo.

2.2.2. Teoría de la distribución de la renta: rechazo a la doctrina del fondo de salarios

Salarios: inicialmente, Mill acepta la doctrina del fondo de salarios, pero se retracta en 1869, en su reseña del libro de W. T. Thornton. La doctrina se usaba en contra de los sindicatos, porque, si el salario solo dependía del fondo de salarios y de la población, la negociación colectiva no podía elevarlo. Persuadido por Thornton de que el fondo no está predeterminado, Mill admite que unos sindicatos que sí dispongan de poder suficiente lograrán elevar los salarios reduciendo el beneficio empresarial en lugar del empleo, y esto, a su vez, puede precipitar la llegada del estado estacionario.

Beneficios e interés: Mill se mantiene fiel al modelo ricardiano básico, que predecía descensos de los beneficios. En general, también defiende la ley de Say para refutar a los subconsumistas, si bien reconoce que puede haber un exceso de oferta en algunos bienes cuando el mercado reacciona ante cambios en las condiciones de oferta y demanda. Ahora bien, afirmó que era ilógico llevar este análisis al conjunto de la economía y concluir que podía haber permanentemente un exceso de oferta de todos los bienes[34].

[32] El único grupo de bienes que dejó sin analizar es el de aquellos con costes decrecientes y una curva de oferta a largo plazo de pendiente negativa.

[33] Recurrió asimismo a este tipo de bienes para examinar el monopolio, en los que el monopolista es capaz de restringir la oferta de forma artificial.

[34] En su ensayo de 1844, Mill distinguió tres tipos de economía y mostró con claridad cuándo puede darse un exceso de oferta agregada. En una economía de trueque (i), ofrecer bienes presupone la existencia de una demanda de bienes y, por tanto, no hay excesos de oferta, y lo mismo ocurre cuando el dinero solo es medio de cambio (ii). Si el dinero tiene más funciones (iii), como la de depósito de valor, quien vende no tiene por qué acudir de inmediato al mercado a comprar. Por ello, aunque su venta cree poder adquisitivo suficiente para el pleno empleo, ese poder adquisitivo puede quedar sin ejercer en el periodo actual y provocar un exceso general de oferta.

2.2.3. Teoría del crecimiento: el estado estacionario "feliz"

Mill no discute que la economía tienda al estado estacionario ricardiano, pero separa dos clases de leyes para matizar esa conclusión:

- *Producción*: son tan inmutables como las de la física.

- *Distribución*: dependen de decisiones humanas y de las instituciones, y por eso pueden reformarse. Como el estado estacionario se puede modificar actuando sobre ellas, deja de ser un destino fijo e impuesto desde fuera de la sociedad.

En concreto, Mill abogaba por educar a la clase trabajadora para que limitara su natalidad, lo que haría posible un salario por encima del de subsistencia. Sostenía, además, que un salario superior a ese nivel no implicaba necesariamente una caída de la inversión, puesto que los trabajadores podían dedicar su ahorro a acumular capital.

Así, Mill cree que el estado estacionario puede ser una situación amable[35], y no forzosamente sombría, pues "una condición estacionaria del capital y de la población no implica un estado estacionario del progreso humano".

2.2.4. Teoría del comercio: la teoría de las demandas recíprocas

John Stuart Mill (1844, 1848) definió con gran precisión el concepto de relación real de intercambio (RRI) y explicó sus causas con la **teoría de las demandas recíprocas**.

Idea: según esta teoría, la RRI depende de la fuerza relativa de la demanda de productos importados de cada país, puesto que esa demanda, a su vez, determina la oferta de productos de exportación.

Conclusiones:

- *Aspectos positivos*: Mill identificó el mecanismo (la demanda recíproca) que fija la RRI dentro de los límites ricardianos.

- *Limitaciones*: el propio Mill advirtió en la tercera edición (1852) que varias relaciones de intercambio pueden satisfacer la ecuación de la demanda internacional, indeterminación que aclararon las curvas de oferta de Marshall y Edgeworth.

> La teoría de las demandas recíprocas se desarrolla en el tema B05, apartado 1.3, como antecedente de la teoría neoclásica del comercio internacional.

Excepción al libre comercio: Mill admite, como único caso en que la protección es defendible por razones puramente económicas, los aranceles temporales con los que un país joven aclimata una industria extranjera adecuada a sus circunstancias. Ese argumento de la *industria naciente*, con el criterio de Mill, se desarrolla en el tema B07, apartado 1.2.2.

[35] Mill lo concibe como un estado pacífico e ilustrado, en el que el tiempo libre puede dedicarse a reforzar los vínculos sociales y a cultivar la propia inteligencia.

2.2.5. Teoría monetaria y bancaria

Debate bullonista: las aportaciones de Mill a la teoría monetaria se producen en el contexto de la prolongación del debate bullonista. Las crisis financieras de 1825 y 1836-1839 abrieron en Inglaterra un debate, zanjado con la Ley Bancaria de Peel (1844), entre:

▪ *Escuela monetaria* (*currency school*): heredera del bullonismo ricardiano y dirigida por Lord Overstone, que defendía que debía existir un máximo a la cantidad de papel moneda que emitía el Banco de Inglaterra y que esta debía tener una contrapartida en oro.

▪ *Escuela bancaria* (*banking school*): dirigida por Thomas Tooke y a la que se adhirió J. S. Mill, la cual defendía que, mientras se mantuviera la convertibilidad en oro, no había que imponer límites cuantitativos a la emisión de papel moneda, pues la doctrina de los billetes reales (y la "ley del reflujo") impediría la sobreemisión.

Con todo, Mill matizó su adhesión. Coincidía con Tooke y Fullarton en que los billetes apenas influyen en los precios al comienzo de la especulación, pero admitía que pueden alimentarla en su fase avanzada y que la Ley de 1844 obliga a frenar antes el crédito. Concluyó, sin embargo, que, al impedir al Banco de Inglaterra prestar durante el pánico, la ley "agrava sustancialmente la intensidad de las crisis comerciales".

Ciclos de crédito: para matizar la ley de Say, introdujo el dinero en el análisis de la producción general, lo cual permitió mejorar considerablemente los argumentos clásicos. Al abordar estas cuestiones, Mill recuperó para la teoría clásica el análisis monetario de H. Thornton (1802), explicó las crisis comerciales por la especulación y el crédito y demostró que, cuando se introduce el crédito, puede darse un exceso general de oferta de bienes[36]. No obstante, Mill afirma que este será breve y seguido del pleno empleo cuando varíen los precios.

2.2.6. Implicaciones de política económica: la economía del bienestar y la imposición

▪ *Intervención pública*: Mill la justifica en las excepciones que admite al principio del *laissez faire*, muchas de las cuales se llaman hoy fallos de mercado. Para ordenar lo que corresponde al Estado, separa dos grupos de funciones:

- *Funciones necesarias* (p. ej. justicia o defensa): el Estado no puede renunciar a ellas, y marcan el mínimo de su intervención.
- *Funciones facultativas o discutibles* (p. ej. regulación del trabajo infantil y determinadas prestaciones): Mill defiende que existan mientras no estorben la iniciativa privada, y en ellas se adivina un primer esbozo del Estado del bienestar.

▪ *Hacienda pública*: la capacidad de pago, entendida como igualdad de sacrificio, y la exención de un mínimo de renta son los dos principios con los que Mill completa en materia tributaria los cánones de Smith (apartado 1.1.6). Sobre la renta que excede ese mínimo propone un tipo proporcional, que da lugar a una progresividad moderada. La escala progresiva de tipos le parece, en cambio, un impuesto sobre la laboriosidad y el ahorro, y solo la acepta en el impuesto sobre las herencias.

[36] Para Mill, ese exceso general de oferta no se debe a una superproducción entendida al modo malthusiano, como superabundancia general, sino a un cambio en las expectativas de los empresarios.

2.2.7. Conclusiones

Aspectos positivos: suele verse en Mill al clásico que más atención prestó a la vertiente social de la economía. Sus aportaciones al análisis de la intervención pública fueron pioneras y han ejercido una gran influencia que llega hasta hoy.

Limitaciones: su eclecticismo dio lugar a una síntesis sin un principio unificador, y su teoría del valor basada en el coste de producción, que él daba por acabada, quedó superada por la revolución marginalista de la década de 1870.

2.3. Conclusiones

Senior y Mill mantienen una teoría del valor basada en el coste de producción y corrigen el sistema clásico en varios puntos. Senior separa la ciencia positiva del consejo político y explica el interés como remuneración de la abstinencia, con lo que distingue el beneficio del interés y anticipa las teorías de la preferencia temporal. Mill da a la demanda un papel en los precios según cómo responda la oferta de cada clase de bienes, y su abandono del fondo de salarios permite a los sindicatos elevar el salario a costa del beneficio.

La separación entre las leyes de la producción, fijas, y las de la distribución, reformables, lleva a Mill a ver en el estado estacionario una situación compatible con el progreso humano. Mill admite, además, que el Estado corrija los defectos de mercado y que el crédito provoque excesos generales de oferta, aunque breves.

3. Las críticas al pensamiento clásico: Karl Marx (*El capital*, 1867)

Formación: doctor en filosofía (Jena, 1841), periodista y activista revolucionario, desde 1849 vivió exiliado en Londres, donde estudió a los clásicos en la biblioteca del Museo Británico.

Importancia: Marx abogó por una revolución fundamental de la economía y la sociedad, y su obra está en el origen de múltiples corrientes de pensamiento económico, político y filosófico. En el plano económico suele situársele al final de la escuela clásica, porque el valor-trabajo sigue siendo el punto de partida de su análisis. *El capital* (1867) es su obra de mayor relevancia para el pensamiento económico. Marx solo publicó el libro I. Por su parte, Engels editó los libros II (1885) y III (1894) a partir de los manuscritos de Marx.

Metodología: Marx emplea un método dialéctico e histórico, que trata el capitalismo como un modo de producción transitorio. En *El capital*, la exposición avanza de lo abstracto (la mercancía, el valor) a lo concreto (los precios, las crisis).

3.1. Rasgos generales y teoría de la historia: el materialismo histórico

El pensamiento de Marx se distingue de la escuela clásica en dos aspectos principales:

- *Dualismo de clases*: los clásicos distinguían tres clases (trabajadores, capitalistas y terratenientes). Marx, en cambio, reduce el esquema a dos, a las que vincula la relación de producción. De un lado están los capitalistas, dueños de los medios de producción, y enfrente el proletariado, que solo tiene su fuerza de trabajo para vender en el mercado. Aunque Marx conserva en su obra a los terratenientes como tercera clase, el conflicto central del capitalismo es para él el que enfrenta a esas dos.

- *Relaciones de poder*: para los clásicos ortodoxos, la economía obedece a leyes "naturales", de validez universal, que ningún actor puede controlar. Marx replica, desde el materialismo histórico, que esas leyes cambian con los procesos históricos y que hay relaciones de poder que desembocan en la lucha de clases.

El análisis del capitalismo de Marx y, en general, su teoría económica son una aplicación de su teoría de la historia (el **materialismo histórico**) a la economía capitalista de su época. Esta teoría de la historia se apoya, a su vez, en la dialéctica de Georg Wilhelm Friedrich Hegel, exponente del idealismo alemán, si bien Marx invierte la dialéctica idealista hegeliana para convertirla en una dialéctica materialista. El objetivo último de su análisis es poner al descubierto las leyes de la dinámica del capitalismo.

Modos de producción: Marx expone su teoría de la historia en el prefacio de la *Contribución a la crítica de la economía política* (1859) y en el *Manifiesto comunista* (1848). En estas obras desarrolla el concepto de *modo de producción*, central en su teoría económica.

La idea principal es que cada época histórica se corresponde con un modo de producción distinto, entendido como la manera en que una sociedad organiza su producción y su vida material. El prefacio de 1859 enumera los modos asiático, antiguo (esclavista), feudal y burgués moderno, y el socialismo es para Marx el que sucederá al capitalismo. Además, Marx sostiene que todos los modos de producción (salvo las sociedades sin clases) llevan aparejado, de manera ineludible, un conflicto de intereses entre dos clases antagónicas (amos y esclavos, señores y siervos, capitalistas y proletarios, según la época).

Todo modo de producción se compone de dos elementos:

- *Fuerzas productivas*: reflejan la tecnología de producción de la sociedad en la fabricación de bienes materiales (p. ej. cualificación de los trabajadores, herramientas o bienes de capital), y constituyen una fuerza dinámica.

- *Relaciones de producción*: son las "reglas del juego" y vienen determinadas históricamente. Reflejan las relaciones de propiedad y sociales, determinan cómo se reparte la producción, llevan consigo un conflicto permanente y constituyen una fuerza estática.

A su vez, de las relaciones de producción se derivan varios conceptos relevantes en la teoría marxista:

- *Estructura económica (base)*: el conjunto que forman las relaciones de producción existentes en una economía.

- *Superestructura jurídica y política*: descansa sobre la estructura económica y sirve para conservar sin cambios las relaciones de producción, es decir, el *statu quo*. La encarnan el Estado y la ideología que le da legitimidad[37].

- *Fetichismo de la mercancía*: fenómeno por el cual las relaciones sociales entre los productores aparecen como relaciones entre cosas (las mercancías). De este modo, el valor se percibe como una propiedad objetiva e intrínseca de las mercancías, ocultando que es la cristalización del trabajo social empleado en producirlas.

- *Alienación*: en los *Manuscritos económico-filosóficos* de 1844, Marx extiende el análisis más allá de lo económico. Bajo el capitalismo, el trabajador se aliena del producto de su trabajo, del propio proceso productivo, de su naturaleza como especie y de los demás hombres, porque tanto el producto como la propia actividad de trabajar pertenecen a otro, el capitalista[38].

La dialéctica materialista: Hegel sostenía que el Espíritu (*Geist*) se despliega en la historia mediante un proceso dialéctico. En la versión simplificada habitual, un estadio inicial o *tesis* genera su contrario o *antítesis*. De la confrontación entre tesis y antítesis surge un estadio superior o *síntesis*, que se convierte a su vez en una nueva tesis. La tríada tesis-antítesis-síntesis procede de Fichte y de los divulgadores de Hegel más que de la obra de este.

Marx no emplea esa tríada, que ridiculiza en *Miseria de la filosofía* (1847), pero conserva la dialéctica y elimina su carácter idealista, de modo que las condiciones materiales, y no el Espíritu, mueven la historia. En la lectura didáctica habitual se identifican las relaciones de producción (estáticas) con la tesis de un estadio social y las fuerzas productivas (dinámicas) con la antítesis. El proceso tiene tres fases:

1. *Armonía*: al comienzo de cada periodo histórico, las fuerzas productivas y las relaciones de producción se corresponden.

2. *Contradicción*: con el tiempo, el desarrollo de las fuerzas productivas choca con las relaciones de producción existentes, que se convierten en una traba, y la contradicción se manifiesta en una lucha de clases.

3. *Síntesis*: la revolución social resuelve la contradicción con un nuevo conjunto de relaciones de producción.

[37] Por superestructura social se entiende la religión, la filosofía, la jurisprudencia, el arte, la música, la literatura y las demás manifestaciones culturales que la sociedad acepta.

[38] Los *Manuscritos* son anteriores al concepto de *fuerza de trabajo*, que Marx elabora a finales de la década de 1850.

Las sociedades pasan así por distintos estadios. En particular, la lucha de clases en el capitalismo conduce al socialismo, que abole la propiedad privada de los medios de producción bajo la dictadura del proletariado. Extinguidas las clases que aún subsisten en esa transición, el socialismo desemboca en el comunismo, una sociedad sin clases ni Estado en la que termina la lucha de clases[39].

3.2. Teoría del valor

Los clásicos buscaban en el valor la explicación de los precios relativos de las mercancías. Marx, en cambio, lo emplea sobre todo para explicar el origen del beneficio y el carácter social del trabajo en una economía mercantil, y deja los precios para el libro III de *El capital*, donde los valores se transforman en precios de producción. Su teoría del valor parte de la teoría ricardiana del valor-trabajo y de una distinción entre productos (con valor de uso) y mercancías (con valor de uso y valor de cambio, producidas para el intercambio). Las mercancías existen antes del capitalismo, pero este generaliza la producción mercantil y convierte en mercancía la propia fuerza de trabajo.

El doble carácter del trabajo: Marx busca el elemento común que dota a las mercancías de su valor de cambio y concluye que el *tiempo de trabajo socialmente necesario* para producirlas determina su valor y, en primera aproximación, sus precios relativos[40]. Para ello distingue el doble carácter del trabajo:

- *Trabajo concreto*: el trabajo útil específico (tejer, forjar, etc.), cualitativamente distinto en cada actividad, que produce el *valor de uso* del bien.

- *Trabajo abstracto*: el trabajo humano considerado como gasto homogéneo de energía, medido en tiempo de trabajo socialmente necesario[41], en el que el trabajo cualificado cuenta como un múltiplo del trabajo simple. Este trabajo produce el *valor*, cuya forma de manifestación es el valor de cambio, de modo que los valores relativos reflejan la cantidad de este trabajo abstracto necesaria para producir cada bien.

Composición del valor: capital constante, capital variable y plusvalía: el valor de una mercancía, $W = c + v + s$, se descompone en tres elementos:

- *Capital constante (c)* o *trabajo muerto* (*dead labour*): valor de los medios de producción consumidos (las materias primas y, en la maquinaria, solo su desgaste) que se transfiere al producto. Se denomina "constante" porque no crea valor nuevo, sino que únicamente transmite su valor preexistente. A diferencia de la distinción de Smith entre capital fijo y circulante, agrupa medios de producción de ambos tipos.

- *Capital variable (v)*: valor de la fuerza de trabajo, es decir, los salarios. Se denomina "variable" porque la fuerza de trabajo crea más valor del que cuesta. Corresponde a la parte del *trabajo vivo* (*living labour*) que reproduce el valor de la propia fuerza de trabajo.

[39] En la *Crítica del programa de Gotha* (1875), Marx distingue una primera fase del comunismo, en la que cada uno recibe según su trabajo, y una fase superior regida por el principio "de cada cual según sus capacidades, a cada cual según sus necesidades". Lenin llamó socialismo a la primera.

[40] Marx también examinó el valor de uso o utilidad (total) como elemento común, pero rechazó esta posibilidad.

[41] Tiempo de trabajo socialmente necesario: el que requiere producir un valor de uso en las condiciones de producción normales de la sociedad y con el grado medio de destreza e intensidad del trabajo.

- *Plusvalía o plusvalor (s)* (*Mehrwert*): el valor adicional creado por el trabajo vivo por encima de v, del que se apropia el capitalista, es decir, la diferencia entre el valor de la mercancía y su *precio de coste* (*Kostpreis*), $c + v$, de modo que $s = W - (c + v)$[42].

El mensaje de Marx es, por tanto, que cualquier economía capitalista produce más valor del necesario para cubrir todos los costes sociales reales de producción[43]. No obstante, convertir estos valores en precios de producción, con una tasa de ganancia uniforme entre industrias, plantea el problema de la transformación.

3.3. Teoría de la distribución de la renta: plusvalía y explotación

En la circulación simple de mercancías se vende para comprar, M-D-M, y el dinero solo es un intermediario. El capital recorre el circuito inverso, D-M-D′, en el que se compra para vender más caro, con D′ > D, y Marx se pregunta de dónde sale el incremento si las mercancías se intercambian por su valor. La respuesta está en la distinción entre *trabajo* y *fuerza de trabajo*, la capacidad de trabajar, que el trabajador vende al capitalista por su valor, es decir, por el tiempo de trabajo necesario para producir sus bienes de subsistencia. Esos bienes no se reducen al mínimo fisiológico, porque el valor de la fuerza de trabajo contiene, según Marx, un "elemento histórico y moral" que depende de las costumbres y del nivel de vida de cada país. La jornada, sin embargo, es más larga que ese tiempo. En una parte de ella (el *trabajo necesario*) el trabajador reproduce el valor de su salario, v, y en el resto (el *plustrabajo*) crea la plusvalía, s. La explotación no requiere, por tanto, ningún intercambio desigual. El capitalista puede aumentar la plusvalía de dos maneras:

- *Plusvalía absoluta*: alarga la jornada o intensifica el trabajo.

- *Plusvalía relativa*: eleva la productividad en los sectores que producen bienes de subsistencia, lo que abarata la fuerza de trabajo y reduce el trabajo necesario.

Como los medios de producción están en poder de los capitalistas, el trabajo no percibe la plusvalía que genera, y por eso, según Marx, el capitalismo genera la alienación y la explotación del trabajador. En *El capital*, la plusvalía se reparte entre el beneficio industrial, el interés y la renta de la tierra.

3.4. Teoría del crecimiento y de los ciclos: las leyes del movimiento capitalista

Según Marx, la dinámica capitalista obedece a cinco leyes o tendencias generales, nacidas todas de la tensión entre unas fuerzas productivas cambiantes y unas relaciones de producción fijas.

Ley de la tendencia decreciente de la tasa de ganancia: como la plusvalía solo nace del trabajo, el capitalista que busca la mayor posible tratará de explotar ese trabajo al máximo. La ley se analiza con tres ratios:

- *Tasa de plusvalía o de explotación*: $s_v = s/v$.

- *Composición orgánica del capital*: Marx la define como c/v, y aquí, por comodidad, se expresa como $\theta = c/(c + v)$[44].

[42] El *precio de coste* puede entenderse como el coste de producción desde el punto de vista del capitalista, lo que a él le cuesta producir la mercancía. Marx emplea un término propio porque el coste de producción de los clásicos solía incluir también el beneficio ordinario.

[43] La plusvalía marxiana guarda así semejanza con el producto neto de los fisiócratas.

[44] Parte de la literatura (Sweezy, 1942; Blaug, 1997, cap. 7) la expresa como $c/(c + v)$. En ambos casos mide el grado de mecanización de la producción. Como $v/(c + v) = 1 - \theta$, se tiene $r = \frac{s}{v} \cdot \frac{v}{c+v} = s_v\,(1 - \theta)$.

- *Tasa de ganancia*: $r = s/(c+v)$.

Operando, las tres se relacionan según:

$$r = \frac{s}{c+v} = s_v\,(1-\theta)$$

Dos motivos pueden hacer caer la tasa de ganancia. El primero, la acumulación de capital, tiene un efecto distinto según se mire desde la empresa o desde el conjunto. Para el capitalista individual que mecaniza antes que sus competidores, la acumulación es rentable, porque ese capitalista produce por debajo del valor social de la mercancía y, mientras no le imiten, obtiene una plusvalía extraordinaria. Cuando todos los capitalistas mecanizan, en cambio, sube la composición orgánica media del capital. Como la plusvalía solo la crea el trabajo vivo, la tasa de ganancia baja salvo que la tasa de plusvalía crezca lo bastante para compensarlo:

$$\uparrow c \;\Rightarrow \uparrow \theta = \frac{c}{c+v} \;\Rightarrow \downarrow r = s_v\,(1-\theta)$$

La ley se deriva con una tasa de plusvalía constante. El segundo motivo, la presión al alza de los salarios, que recorta la plusvalía y, con ella, la tasa de ganancia, pertenece en cambio al ciclo de la distribución y no a la ley ($\uparrow v \Rightarrow \downarrow s \Rightarrow \downarrow s_v \Rightarrow \downarrow r$).

Al igual que Smith, Ricardo y Mill, Marx concluye que la tasa de beneficio tiende a descender, si bien por un mecanismo distinto, el aumento de la composición orgánica del capital, y no por los rendimientos decrecientes (Ricardo) ni por la competencia (Smith).

Ley de concentración: aumento del poder de mercado: la gran empresa, más especializada y con una capacidad productiva muy superior, va desplazando a los pequeños talleres, porque el afán de beneficio empuja a mecanizar.

Ese aumento de la capacidad productiva puede desembocar en sobreproducción y en una caída de precios que solo resisten los productores más eficientes. A la larga, el capital industrial se reúne en menos empresas y el control de la economía pasa a un grupo cada vez más reducido.

Ley del ejército industrial de reserva: existencia de desempleo involuntario: Marx rechazó la teoría malthusiana de la población y la sustituyó por una ley de población propia del capitalismo. Según ella, la mecanización deja sin empleo a trabajadores desplazados por las máquinas (parados tecnológicos) o por las crisis (cíclicos). Esa *sobrepoblación relativa* es el "ejército industrial de reserva", que presiona los salarios a la baja y los contiene dentro de límites compatibles con la acumulación.

De esta manera, según Marx puede existir desempleo involuntario, lo cual equivale a rechazar la ley de Say en la lectura que hará Keynes (1936), para quien esta implica el pleno empleo de recursos.

Ley de la miseria creciente del proletariado: aumento de la desigualdad: cuanto más numeroso es el ejército de reserva, peor es la situación del proletariado. A ello se suma que los capitalistas, para frenar el descenso de sus beneficios, endurecen las condiciones de trabajo (jornadas más largas, salarios más bajos, empleo de mujeres y niños…).

Ley de las crisis: ausencia de estabilidad económica: Marx no ofrece una teoría única de las crisis. Describe varios mecanismos (la caída de la tasa de ganancia, la desproporción entre los sectores de los esquemas de reproducción y el subconsumo de los trabajadores), todos con la acumulación de capital (la inversión) como motor común. Los *esquemas de reproducción* del libro II dividen la economía en dos sectores, el I, que produce medios de producción, y el II, que produce bienes de consumo. La reproducción simple exige que el sector I venda al II tantos

medios de producción como bienes de consumo compran al II los trabajadores y los capitalistas del I, $v_1 + s_1 = c_2$, y nada asegura que la inversión descentralizada respete esa proporción. La crisis es posible porque en una economía monetaria la venta y la compra se separan (M-D-M), por lo que Marx rechaza la ley de Say.

El ejército de reserva genera, además, un ciclo de la distribución. Cuando los salarios son elevados (p. ej. por una disminución del desempleo), la plusvalía disminuye, y con ella la tasa de ganancia y, por tanto, la acumulación de capital, mientras que, cuando los salarios son bajos, se produce la situación opuesta.

Implicación principal de las leyes del movimiento capitalista: las leyes anuncian el fin del capitalismo. La burguesía se apropia de la plusvalía que crea el trabajo asalariado, y la acumulación concentra el capital en cada vez menos manos mientras crece la miseria del proletariado. Ese mismo proceso agrupa y organiza a una clase obrera cada vez más numerosa, cuya rebelión termina con la expropiación de los expropiadores y da paso al socialismo, un nuevo modo de producción. En *El capital*, el tránsito depende, por tanto, de la acción de la clase obrera y no de un derrumbe automático del sistema, tesis que defenderán autores posteriores (Luxemburg, 1913; Grossmann, 1929).

Conclusiones:

- *Aspectos positivos*: Marx trata el ciclo y el cambio técnico como rasgos endógenos del capitalismo, una visión evolutiva que Schumpeter (1942) le reconoció, y su interés por el gasto en inversión ha sido compartido por muchos macroeconomistas modernos.

- *Limitaciones*: cada una de las leyes presenta objeciones:

  - *Tasa de ganancia*: la principal objeción es que Marx da poca importancia a un efecto de signo contrario. La acumulación de capital puede hacer más productivo el trabajo y elevar así la tasa de plusvalía, y puede abaratar también los medios de producción y frenar la subida de la composición orgánica en valor (el propio Marx admitió estas "tendencias contrarrestantes"). Por eso no es seguro que una mayor dotación de capital rebaje la tasa de ganancia[45].

  - *Ley de concentración*: se ha cumplido en parte (De Loecker, Eeckhout y Unger, 2020, documentan el aumento reciente de los márgenes), pero la competencia internacional y la deslocalización en cadenas globales de valor, que Marx no previó, la han matizado.

  - *Ley del ejército industrial de reserva*: la objeción habitual es que, con mercados competitivos, el salario debería ajustarse hasta absorber el exceso de oferta de trabajo. En Marx, sin embargo, ese ajuste es el propio ciclo de la distribución, y el cambio técnico que ahorra trabajo recrea el ejército de reserva. La objeción de fondo es empírica, porque la mecanización no ha generado un paro creciente a largo plazo.

  - *Ley de la miseria creciente*: si la miseria se entiende en términos de renta real o de participación de los trabajadores en la renta nacional, la tendencia no se ha cumplido. Por ejemplo, en Gran Bretaña, según Allen (2009), los salarios reales, estancados hasta la década de 1840, crecieron después al ritmo de la productividad. La participación del trabajo en la renta, que

[45] Okishio (1961) demostró que, con el salario real constante, un cambio técnico que los capitalistas adoptan porque reduce los costes a los precios vigentes no puede reducir la tasa de ganancia de equilibrio. La caída exige, por tanto, que suba el salario real, un efecto que pertenece al ciclo de la distribución y no a la ley.

había caído hasta mediados del s. XIX, volvió hacia 1900 a su nivel de finales del s. XVIII. En cambio, si se entiende en términos no económicos, resulta difícilmente medible[46].

- *Ley de las crisis*: Marx no llega a desarrollar una teoría del ciclo rigurosa, sino que esta aparece dentro de descripciones más generales de su análisis histórico.

3.5. Implicaciones de política económica: el *Manifiesto comunista* (1848)

El *Manifiesto comunista* (1848), escrito con Engels casi dos décadas antes de *El capital*, es el texto en que Marx y Engels concretan un programa de política económica. El texto presenta ese programa como un conjunto de medidas de transición, con las que el proletariado, una vez conquistado el poder político, arrebataría gradualmente el capital a la burguesía y centralizaría los medios de producción en manos del Estado. Admite, además, que las medidas variarían según el país, y propone diez para los más avanzados.

El núcleo del programa ataca la propiedad privada, origen de la apropiación de la plusvalía en el análisis posterior de Marx. Así, propone abolir la propiedad de la tierra y destinar su renta a fines públicos, suprimir el derecho de herencia y confiscar los bienes de emigrados y rebeldes. A ello se suman la centralización del crédito en un banco nacional con capital del Estado y monopolio exclusivo, la de los transportes y las comunicaciones, y la extensión de las fábricas y de los medios de producción estatales, junto con el cultivo de las tierras baldías conforme a un plan común.

Otras medidas organizan el trabajo y la sociedad: un impuesto sobre la renta fuertemente progresivo, la obligación de trabajar para todos, con ejércitos industriales sobre todo para la agricultura, la combinación de la agricultura y la industria para borrar poco a poco la diferencia entre el campo y la ciudad, y la educación pública y gratuita de todos los niños, con la abolición del trabajo infantil en las fábricas "en su forma actual".

La suerte de estas propuestas fue desigual. La propiedad estatal de los medios de producción y la planificación definieron las economías socialistas del siglo XX. Otras, en cambio, se incorporaron a las economías de mercado sin alterar su naturaleza: el impuesto progresivo sobre la renta, la educación pública gratuita, la prohibición del trabajo infantil y un banco central con el monopolio de emisión.

3.6. El fracaso de las economías comunistas y de las teorías del valor-trabajo

El fracaso de las economías comunistas se explica por el problema de la planificación central, y el de las teorías del valor-trabajo, por el problema de la transformación.

3.6.1. ¿Por qué fracasaron las economías comunistas? El problema de la planificación central

Cálculo económico: Mises (1920) sostuvo que, sin propiedad privada de los medios de producción, no hay precios de los bienes de capital y el planificador no puede calcular qué asignación es eficiente. En respuesta, Lange (1936) propuso un socialismo de mercado con precios fijados por ensayo y error, y Hayek (1945) replicó que el problema es de conocimiento, porque la información

[46] Marx defendió de forma reiterada que entre los peores efectos del capitalismo figura el deterioro de esos factores intangibles que se agrupan bajo la denominación de calidad de vida. La especialización y la división del trabajo engendran a un trabajador "lisiado por la repetición, durante toda su vida, de una misma operación trivial y reducido así a un mero fragmento de hombre".

está dispersa y solo los precios la agregan. Hayek (1988) llamó *fatal arrogancia* a la creencia de los planificadores de que disponen de información completa, y ya en 1944 había sostenido que la planificación central acaba con la libertad política.

> El debate sobre el cálculo económico socialista se desarrolla en el tema A03, apartado 1.4.1 (Mises), y en el apartado 1.4.2 del mismo tema (Hayek).

Incentivos: Kornai (1980) añadió que las empresas socialistas operan con una *restricción presupuestaria blanda*, porque el Estado cubre sus pérdidas. Carecen, por tanto, de incentivos para controlar los costes e innovar, y la economía sufre una escasez crónica.

3.6.2. ¿Por qué fracasaron las teorías del valor-trabajo? El problema de la transformación (von Böhm-Bawerk)

La principal limitación de las teorías del valor-trabajo es el llamado **problema de la transformación**, que Engels anunció en 1885, Böhm-Bawerk planteó en 1896 y Bortkiewicz resolvió formalmente en 1907. El problema consiste en que la teoría del valor-trabajo no explica de forma coherente cómo se convierten los valores en precios de producción cuando la composición orgánica del capital difiere entre industrias y la tasa de beneficio es uniforme.

Marx lo abordó aplicando la tasa media de ganancia a $c + v$, con la suma de los precios igual a la de los valores y la de los beneficios igual a la de las plusvalías, pero valoró los insumos a sus valores. Bortkiewicz mostró que, corregido ese error, los precios de producción pueden calcularse, pero ambas igualdades no se cumplen a la vez en general y la tasa de ganancia en precios difiere de la tasa en valores, $s/(c + v)$.

Samuelson (1971) concluyó que valores y precios son dos sistemas alternativos. Pasar de uno a otro equivale a borrar el primero y escribir el segundo, y los valores no añaden nada a la explicación de los precios. Por su parte, la *nueva interpretación* (Duménil, 1980; Foley, 1982) aplica las igualdades de Marx al producto neto y no al bruto, y define el valor de la fuerza de trabajo a partir del salario monetario, no de la cesta de bienes que este compra. Así, el producto neto a precios equivale al trabajo vivo total y los beneficios a la plusvalía, aunque por definición y no como explicación de los precios.

Para la mayoría de los economistas, el problema de la transformación debilitó seriamente el conjunto de la teoría marxista.

Ante esta situación, se produjeron dos tipos de respuesta:

1. *Pensamiento ortodoxo*: la economía académica ya había sustituido la teoría del valor-trabajo por el análisis marginal en la década de 1870 (tema A03). En el problema vio la confirmación de que el valor debía explicarse con la utilidad marginal, la productividad y la oferta y la demanda.

2. *Pensamiento marxista*: quedó relegado a la heterodoxia, y muchos autores pasaron a centrarse en los fenómenos sociales, con una comprensión holística de la economía ajena a la determinación de los precios. La crítica neorricardiana, a partir de Sraffa (1960), mostró además que los precios de producción y la tasa de ganancia se determinan a partir de las condiciones técnicas y del salario real, sin pasar por los valores-trabajo, que resultan así redundantes (Steedman, 1977).

3.7. Conclusiones

Aspectos positivos: la aportación de Marx pone sobre la mesa la distribución de la renta y el conflicto entre clases en un momento de evidentes problemas de equidad en las economías industrializadas. Su legado analítico es, además, duradero. Los esquemas de reproducción son la base de los modelos de crecimiento con dos sectores (Feldman, 1928), el ciclo entre distribución y acumulación lo formalizó Goodwin (1967), y Leontief (1938) vio en Marx una fuente inagotable de observación directa del capitalismo.

Limitaciones: por un lado, las leyes del movimiento capitalista se han cumplido solo en parte, y la revolución no llegó a los países capitalistas avanzados, donde Marx la esperaba. Por otro lado, la teoría del valor-trabajo tropieza con el problema de la transformación y resulta redundante para explicar los precios. Finalmente, el carácter normativo de la explotación es discutido. Marx la presenta como compatible con el intercambio de equivalentes y evita apelar a la justicia (Wood, 1972), pero el término conserva una carga peyorativa que dificulta separar el análisis positivo del juicio de valor[47]. A estas críticas se suman las razones del fracaso de las economías de planificación central.

[47] Desde la economía neoclásica se objeta que un salario bajo no implica "explotación" si el bienestar del trabajador supera su utilidad de reserva, porque puede preferir trabajar para una multinacional con un salario bajo a estar desempleado. En Marx, sin embargo, la explotación es una magnitud, s/v, que no depende de la alternativa del trabajador. Roemer (1982) la redefinió a partir del reparto desigual de la propiedad de los medios de producción, sin recurrir al valor-trabajo.

Conclusiones

La escuela clásica convirtió la economía política en una disciplina sistemática, con una teoría del valor basada en el coste y, en último término, en el trabajo. Smith situó el origen de la riqueza en la división del trabajo y la acumulación de capital, y Malthus añadió la presión de la población y el riesgo de una demanda insuficiente. Ricardo construyó el primer modelo de la distribución, que predice la caída del beneficio y el estado estacionario, y formuló la ventaja comparativa. Mill sintetizó el sistema y lo abrió a la reforma social, y Marx lo convirtió en una crítica del capitalismo. El sesgo de la escuela hacia la oferta dejó, sin embargo, sin resolver la paradoja del valor y, en Marx, el problema de la transformación.

Las ideas clásicas siguen vivas, transformadas, en la teoría posterior. El análisis marginal sustituye la teoría del valor-trabajo y resuelve la paradoja del agua y los diamantes al separar la utilidad total de la marginal (tema A03), y la mano invisible encuentra su versión formal en el primer teorema del bienestar, que precisa en qué condiciones el equilibrio competitivo es eficiente (tema A22). En sentido contrario, la ley de Say y la dicotomía clásica son el blanco de la crítica de Keynes, que vio en la demanda insuficiente de Malthus un antecedente de la demanda efectiva (tema A04). Por otro lado, la ventaja absoluta, la ventaja comparativa y las demandas recíprocas fundan la teoría del comercio internacional (tema B05), y la equivalencia ricardiana reaparece en la nueva macroeconomía clásica como neutralidad de la financiación del gasto público (tema A06).

La teoría del crecimiento unificado (Galor, 2011) explica la larga etapa malthusiana de las economías preindustriales y su salida mediante la transición demográfica. En el terreno de la distribución, Piketty (2014) ha recuperado la cuestión de la distribución entre el capital y el trabajo, con referencias expresas a Ricardo y a Marx. Asimismo, la caída de la participación del trabajo en la renta en la mayoría de los países desde la década de 1980 (Karabarbounis y Neiman, 2014) ha reabierto el debate sobre sus causas. La ventaja comparativa ricardiana, por su parte, es la base de los modelos cuantitativos de comercio actuales (Eaton y Kortum, 2002; Costinot, Donaldson y Komunjer, 2012).

Anexo. Antecedentes: mercantilismo y fisiocracia (ss. XVI-XVIII)

La escuela clásica tiene dos antecedentes. El mercantilismo, que vinculaba el crecimiento al atesoramiento de oro y defendía el proteccionismo, fue objeto de la crítica de Smith. La fisiocracia, con la tabla económica de Quesnay, anticipó el flujo circular de la renta.

A.1. Mercantilismo

El mercantilismo ocupa el periodo que va de 1500 a 1750 y tuvo sus principales focos en Inglaterra y Francia. Entre los autores más destacados de esta corriente figuran Gerard de Malynes y Thomas Mun.

En la visión mercantilista, el crecimiento económico dependía ante todo de factores monetarios y no de los reales, de modo que la política debía buscar la acumulación de oro y riqueza. Esa meta llevaba a dos recomendaciones. La primera era fomentar el aumento de la población, pues, según esta escuela, solo con muchos trabajadores pobres pero laboriosos, pagados con el salario de subsistencia, podía reunirse esa riqueza.

La segunda era el proteccionismo, porque atesorar exige no incurrir en déficits comerciales. Los aranceles tenían que resguardar en primer lugar a la manufactura, el sector que los mercantilistas consideraban más importante.

Por último, en teoría monetaria sobresale Cantillon (1755), que hace de puente entre mercantilismo y fisiocracia. A su juicio, un aumento del dinero en circulación tiende de entrada a subir los precios, aunque el resultado varía según a qué manos llegue el dinero nuevo:

1. *Mayor gasto*: si el nuevo dinero va a parar a quienes más gastan, los precios suben.

2. *Mayor ahorro*: si va a parar a quienes más ahorran (los prestamistas), el mercado de fondos prestables registra un descenso del tipo de interés.

De este análisis, conocido como *efecto Cantillon*, arranca la discusión monetaria posterior sobre la repercusión que tiene el dinero nuevo en precios y tipos de interés.

Principales críticas: sus ideas no forman un sistema, porque los escritos mercantilistas tratan casi siempre cuestiones económicas sueltas y poco conectadas entre sí. A ello se suman los conceptos erróneos que Smith les reprocha y el sesgo de sus propuestas. Sobre este sesgo, Smith denuncia que fueron los productores quienes idearon el sistema mercantil, en defensa de sus intereses y con total olvido de los consumidores.

A.2. Fisiocracia

Los fisiócratas consideraban la "tabla económica" de Quesnay su principal aportación teórica. En ella se representaba, de forma aproximada, cómo fluían las rentas monetarias entre los diversos sectores económicos (la clase propietaria o terratenientes, la clase productiva o agricultores y la clase estéril o artesanos y comerciantes). La tabla recogía también la creación y la circulación del producto neto (excedente agrario) a través de toda la economía.

El modelo del flujo circular de la renta permite ilustrar de modo intuitivo cómo funciona la actividad económica. Sus antecedentes se encuentran ya en el *Tableau économique* (1758) del

fisiócrata francés François Quesnay, si bien el modelo no fue propiamente desarrollado y popularizado hasta el s. XX gracias a las aportaciones de S. Kuznets, R. Stone y F. Knight[48].

1. *Agentes*: se trata de una economía simplificada, cerrada y sin sector público, en la que coexisten dos agentes económicos:

- *Empresas*: fabrican bienes y servicios y alquilan los factores productivos de los hogares (capital y trabajo), con lo que generan rentas en favor de estos.

- *Hogares*: son propietarios de los factores de producción y adquieren los bienes y servicios producidos con la renta que obtienen al alquilarlos.

2. *Triple dimensión*: de esta forma, los agentes desempeñan tres papeles:

- *Productores*: como oferentes de bienes, son, a su vez, demandantes de factores.

- *Oferentes de factores*: son, a su vez, demandantes de bienes.

- *Perceptores de renta*: son, a su vez, agentes de gasto.

La corriente física de bienes, servicios y factores productivos da lugar, así, a una corriente monetaria de renta y gasto en la economía. Esta idea está en la base de la identidad de la renta nacional que es el punto de partida de la teoría macroeconómica, $Y = C + I$, y que tiene una interpretación alternativa desde el punto de vista del ahorro y la inversión, $S = I$.

Por otro lado, los fisiócratas no se limitaron a teorizar acerca de cómo se relacionaban los diversos sectores económicos, sino que también trataron de medir la magnitud de esas relaciones. La fisiocracia anticipó así la tabla *input-output* que el premio nobel Wassily Leontief elaboró en la década de 1930.

[48] Knight fue el primero en representarlo en forma de diagrama (1933).

A3. Los economistas neoclásicos

Los economistas neoclásicos.

Introducción

Introducción común. Los temas A02 a A07 comparten una introducción, que se desarrolla al comienzo del tema **A02**. Recorre la evolución del pensamiento económico, del mercantilismo y la fisiocracia a la escuela clásica, la revolución marginalista, Keynes y la síntesis neoclásica, el monetarismo, la nueva macroeconomía clásica, la nueva economía keynesiana y la nueva síntesis neoclásica.

La **revolución marginalista** designa el cambio de paradigma que, entre 1871 y 1874, sustituye la teoría clásica del valor, fundada en el coste de producción y en el trabajo incorporado, por otra basada en la **utilidad marginal** (la utilidad de la última unidad consumida). El descubrimiento se produce de forma simultánea e independiente en tres medios académicos incomunicados entre sí (Menger en Viena, Jevons en Manchester y Walras en Lausana).

Con ella, el centro del análisis se desplaza de la producción y el reparto del excedente, la pregunta de los clásicos, a la asignación de recursos escasos entre fines alternativos, y se fija el método que la microeconomía conserva desde entonces, basado en agentes que optimizan en el margen y en precios que coordinan sus decisiones. De esa corriente proceden la teoría del consumidor y de la demanda (temas A08 y A09), el equilibrio general (tema A21) y la economía del bienestar de Pareto (tema A22). Su nombre se debe a Thorstein Veblen, que en 1900 llamó "neoclásica" a la economía de Marshall para subrayar su ascendencia clásica.

Este tema ordena el periodo (1871-1936) por esferas geográficas, porque en cada una de ellas el marginalismo adopta un método y un objeto distintos:

- *Austro-alemana*: subjetivismo y razonamiento verbal.
- *Anglosajona*: equilibrio parcial y aparato instrumental y, en Estados Unidos, extensión del principio marginal desde la utilidad hasta el producto (distribución y capital).
- *Franco-italiana*: equilibrio general y sistema de ecuaciones.

Dentro de cada esfera se sigue una misma secuencia, que va de los *antecedentes* (precursores que aplican el razonamiento marginal a problemas concretos sin formar escuela) a la *fundación* (el autor que convierte la utilidad marginal en el principio del que se deriva todo el sistema y funda una escuela), la *sistematización* (los autores que completan la teoría por el lado de la oferta y de la distribución) y la *evolución posterior* (los desarrollos que extienden o corrigen el sistema desde dentro). El corte entre precursor y fundador lo marca el alcance (la generalidad de la teoría y la continuidad institucional de la escuela), no la cronología.

Tres anexos recogen las diferencias entre clásicos y neoclásicos, las críticas al núcleo teórico y la escuela de Estocolmo. Los programas formales y las construcciones gráficas que aquí solo se enuncian se derivan en los temas de microeconomía a los que se remite en cada caso.

1. Esfera austro-alemana

La primera esfera es el mundo germanohablante, donde el pensamiento económico estaba dominado por la escuela historicista alemana, hostil a la teoría abstracta y partidaria del análisis inductivo e histórico. Ese ambiente adverso explica tanto la irrelevancia inicial de los precursores alemanes (von Thünen y, sobre todo, Gossen) como el carácter marcadamente metodológico que adopta la reacción de Carl Menger desde Viena en 1871.

La tradición que Menger inaugura se caracteriza por el subjetivismo del valor, el razonamiento deductivo verbal sin formalización matemática y la atención al tiempo, la incertidumbre y el empresario. Su desarrollo discurre por una línea de discipulado bien definida (Wieser y Böhm-Bawerk y, después, Mises y Hayek) y por dos controversias que fijan su identidad, la *Methodenstreit* con el historicismo y el debate sobre el cálculo económico socialista.

1.1. Antecedentes: proto-marginalistas en la esfera austro-alemana

Antes de Menger, dos autores alemanes habían aplicado el razonamiento marginal a problemas económicos concretos sin llegar a formar escuela. Su recuperación es póstuma y su influencia, indirecta. Von Thünen actúa sobre la teoría de la productividad marginal que después desarrollan Marshall y Clark, y Gossen sobre la teoría de la utilidad de la primera generación marginalista.

1.1.1. Johann Heinrich von Thünen (1826; 1850)

Von Thünen aplicó el análisis marginal por medio del cálculo (optimización) y contribuyó a desarrollar la productividad marginal como base de la explicación de los salarios. Así, sostuvo que el empresario contrata trabajadores hasta que el producto del último iguala su salario, anticipando en medio siglo la teoría marginal de la distribución (apartado 2.5.1).

En su obra (1826; 1850), de influencia explícita sobre Marshall, formuló además el primer modelo espacial de la economía, donde los cultivos se ordenan en anillos concéntricos en torno al mercado según el coste de transporte. El modelo explica así la renta de la tierra por su localización y no solo por su fertilidad como en Ricardo.

1.1.2. Hermann Heinrich Gossen (1854)

Gossen aplicó el análisis marginal en su obra de 1854, ignorada en su época y que Jevons rescató en el prefacio a la segunda edición de su libro (1879). De ese análisis derivó dos enunciados sobre la utilidad marginal, bautizados después como *leyes de Gossen*[1]:

1. *1.ª Ley (utilidad marginal decreciente)*: sostiene que el placer que proporciona cada nuevo acto de consumo es menor que el del anterior, hasta alcanzar la saciedad. La idea, sin embargo, no se debe originalmente a Gossen, ya que autores como Bernoulli o Bentham la habían planteado antes.

2. *2.ª Ley (principio de equimarginalidad)*[2]: para maximizar su utilidad total, los consumidores distribuyen sus compras de forma que la última unidad monetaria que gastan en cualquier bien les reporta la misma utilidad que la última unidad monetaria gastada en cualquier otro.

[1] La tercera (la escasez como prerrequisito del valor económico) se formula con posterioridad. Gossen no habla explícitamente de renta, sino de tiempo.

[2] Formalmente, $\frac{\partial u/\partial x_i}{p_i} = \frac{\partial u/\partial x_j}{p_j}$, $\forall i, j$. Gossen la formula con utilidad cardinal, pero la condición solo depende de cocientes de utilidades marginales y se mantiene con utilidad ordinal.

1.2. Carl Menger (*Principios de economía*, 1871)

Formación: estudió Derecho en Viena y Praga, se doctoró en Cracovia (1867) y fue periodista y funcionario de la oficina de prensa del Gobierno austriaco antes de incorporarse como profesor a la Universidad de Viena (1873).

Importancia: Menger, en 1871, convierte la utilidad marginal en el principio del que se deriva todo el sistema y funda una escuela que le da continuidad, dos cosas que von Thünen y Gossen no hicieron.

Metodología: Menger se apoya en la lógica deductiva para defender sus teorías, sin recurrir al formalismo matemático. Esta posición, sistematizada en sus *Investigaciones sobre el método de las ciencias sociales* (1883), le enfrentó a Gustav von Schmoller y a la joven escuela historicista alemana en la controversia conocida como *Methodenstreit* (tema A01, apartado 2.2.2).

1.2.1. Teoría del valor: demanda, oferta y equilibrio

Entre quienes resolvieron la paradoja del agua y los diamantes que había planteado Smith (A02), Menger fue de los más claros. El agua tiene más valor de uso que los diamantes y, sin embargo, menos valor de cambio.

El origen de la paradoja está en confundir dos magnitudes, la utilidad total de un bien, ligada a su valor de uso, y su utilidad marginal, la que aporta la última unidad consumida. La utilidad marginal, y no la total, fija el valor de cambio. Como la marginal decrece con la cantidad, la última unidad de un bien tan abundante como el agua aporta muy poca utilidad, mientras que la del diamante, escaso, sigue aportando mucha.

1.2.2. Teoría de la producción y la distribución

En su obra, Menger también sentó las bases para el análisis marginal de los factores de oferta al realizar la siguiente clasificación de bienes:

1. *Bienes de orden inferior o de primer orden* (es decir, bienes finales): satisfacen necesidades humanas directamente (p. ej. pan).

2. *Bienes de orden superior (segundo, tercer orden y sucesivos)* (es decir, factores de producción): satisfacen necesidades humanas de forma indirecta, porque contribuyen a producir bienes de primer orden. En el caso del pan, la harina es de segundo orden, el trigo y los molinos son de tercero, y las tierras de cultivo y los aperos, de cuarto.

Menger atribuye a los bienes de orden superior dos propiedades, la complementariedad y la imputación:

- *Complementariedad*: un bien de orden superior solo es un "bien" cuando se combina con otros de orden superior, porque los factores únicamente producen juntos. Basta con que falte uno imprescindible para el proceso para que el resto no le sirva de nada al productor.

- *Imputación*: un bien de orden superior toma su valor del bien de primer orden a cuya producción contribuye. Por tanto, si este deja de ser un bien, los factores que lo producen corren la misma suerte.

Para Menger, en suma, el valor de un bien de primer orden lo fija su utilidad marginal, no el trabajo incorporado ni lo que cuestan los factores (los bienes de orden superior) empleados en producirlo. En ello Menger se separa de los clásicos. El elemento subjetivo pasa así a ser el determinante del valor, idea central de la escuela austriaca.

Menger incorpora además el tiempo y la incertidumbre a la valoración de los bienes. Pasar de los bienes de orden superior a los de primer orden requiere tiempo, y el tiempo obliga a tener en cuenta la incertidumbre sobre las condiciones futuras. Los bienes presentes valen, por tanto, más que los futuros, puesto que no requieren transformación (o ya la han sufrido) y no están sujetos a incertidumbre. Menger prefigura así el concepto de **preferencia temporal**, que desarrollaría explícitamente uno de sus discípulos, Eugen von Böhm-Bawerk.

Para Menger, que los individuos maximicen su utilidad no asegura que los mercados estén en equilibrio, porque, a diferencia de otros pioneros del marginalismo, tiene en cuenta la incertidumbre. Por ello concede un papel central al empresario, el agente que busca información y detecta oportunidades de beneficio, y en esto le sigue en general la escuela austriaca. Sobre esa base, los austriacos posteriores (Mises, 1949; Hayek, 1948; Kirzner, 1973) entenderán la competencia como un proceso dinámico de descubrimiento, frente a la noción estática de Jevons y Walras.

1.2.3. Teoría monetaria y bancaria: el origen del dinero

Menger explica el dinero como una institución que surge de forma espontánea, sin acuerdo previo ni intervención del Estado (en sus *Principios* y en su artículo de 1892). En el trueque, cada agente acepta a cambio de lo que ofrece los bienes que pueden revenderse con más facilidad, aunque no los necesite. La práctica, imitada por los demás, acaba por convertir en medio general de cambio a la mercancía más vendible (*Absatzfähigkeit*), normalmente los metales preciosos. Mises partió de esta explicación para fundar el valor del dinero en su teorema de la regresión.

1.2.4. Conclusiones

Aspectos positivos: su teoría subjetiva del valor y su clasificación de los bienes por órdenes, base de la imputación, fundan la escuela austriaca, que sus discípulos continúan desde la cátedra de Viena.

Limitaciones: Menger aplica tímidamente el análisis marginal al lado de la oferta, pero, como otros marginalistas de primera generación, se centra sobre todo en el lado de la demanda, lo que no permite determinar completamente los precios[3].

1.3. Sistematización: Wieser y Böhm-Bawerk

Menger había explicado el valor de los bienes finales por la utilidad marginal y mostrado que de él deriva el de los factores, pero no cómo repartirlo entre ellos ni de dónde procede el rendimiento del capital. Sus discípulos directos completan el sistema, Wieser con la imputación del valor a los factores y el coste de oportunidad, y Böhm-Bawerk con la teoría del capital y el interés.

[3] En el intercambio aislado, Menger solo fija un intervalo de precios posibles, entre las valoraciones de las dos partes. Böhm-Bawerk (1889) lo estrecha con su mercado de caballos, en el que el precio queda acotado por dos *parejas marginales*, la del último comprador que adquiere y el primer vendedor excluido, y la del último vendedor que vende y el primer comprador excluido.

1.3.1. Teoría del coste de oportunidad y teoría de la imputación (Wieser, 1884; 1889)

Wieser utilizó el concepto de utilidad marginal[4] para prolongar un paso más la cadena causal de Menger, que va de la utilidad marginal al precio del bien de consumo final, e imputar ese valor a los bienes de orden superior (los factores) empleados en fabricarlo.

Wieser, en su estudio del precio de los factores, fue el primer autor en introducir explícitamente la noción de **coste de oportunidad**, el valor de la mejor alternativa a la que se renuncia al destinar un factor a un uso determinado. La idea había sido apuntada por autores anteriores (von Thünen, 1826; Mill, 1848), pero Wieser le dio un desarrollo riguroso y sistemático. De ahí que la determinación del valor de un factor por su mejor empleo alternativo se conozca como ley de Wieser[5].

1.3.2. El problema del interés: las teorías del capital y el interés (Böhm-Bawerk, 1889)

Origen: los austriacos, y en particular Böhm-Bawerk, conciben el capital no como un factor originario más (para Böhm-Bawerk los únicos factores originarios son la tierra y el trabajo), sino como trabajo y tierra acumulados que permiten *métodos de producción indirectos* o *roundabout*. El capital es, así, producción que toma tiempo, y su empleo más prolongado eleva el producto obtenido.

Esta concepción permite abordar el llamado *problema del interés*. Conforme a la teoría de la imputación, el valor de los medios de producción deriva íntegramente del valor del producto. ¿Por qué, entonces, el valor del producto supera de forma sistemática el de los medios de producción empleados y deja un excedente que percibe el capital en forma de interés?

Idea: Böhm-Bawerk (1889) explica que el tipo de interés es positivo porque los bienes presentes valen más que los bienes futuros de igual clase y cantidad.

Razonamiento: Böhm-Bawerk aduce tres razones (los "tres motivos"):

- *Circunstancias presentes y futuras*: quien espera estar mejor provisto en el futuro que hoy (p. ej. en una economía en crecimiento) valora más los bienes presentes, que para él son relativamente más escasos.

- *Miopía de los agentes* (falta de imaginación, incertidumbre): lleva a infravalorar sistemáticamente las necesidades futuras.

- *Superioridad técnica de los bienes presentes*: permiten iniciar de inmediato métodos de producción más indirectos (*roundabout*), que son más productivos (p. ej. pescar con red, fabricada renunciando a pescar hoy, frente a pescar con la mano). Sin una remuneración del capital, ningún empresario se decidiría por esos métodos indirectos.

El interés es, así, el *agio* o prima que iguala el valor de los bienes presentes y futuros. Böhm-Bawerk daba primacía al tercer motivo, que a su juicio bastaba por sí solo para que el interés fuese positivo.

Extensiones: Böhm-Bawerk empleó su teoría del interés y del capital (y, en general, el marginalismo) para criticar la teoría marxista del valor-trabajo y la idea de explotación (Anexo A.1), a lo que replicó Hilferding (Anexo A.2).

[4] De hecho, fue quien acuñó el término, ya que Menger no había empleado la expresión utilidad marginal.
[5] El término inglés *opportunity cost* se debe, no obstante, a David I. Green (1894).

Conclusiones:

- *Aspectos positivos*: ofrece una explicación sistemática de por qué el tipo de interés es positivo, al combinar la preferencia temporal con la productividad de los métodos indirectos. Fisher la tomó como base de la teoría del interés aceptada hoy (apartado 2.5.2), y la concepción de Böhm-Bawerk del capital como producción que toma tiempo inspira la teoría austriaca del ciclo.

- *Limitaciones*: la pretensión de independencia del tercer motivo fue refutada por Irving Fisher, quien mostró que la productividad de los métodos *roundabout* no es independiente de las dos primeras razones.

1.4. Evolución posterior: Mises y Hayek

Con la tercera generación cambia el objeto. Mientras Menger, Wieser y Böhm-Bawerk trabajan sobre el valor, el precio y la distribución, Mises y Hayek desplazan el análisis hacia la moneda, el ciclo, la crítica de la planificación y, en último término, la metodología y la teoría del conocimiento.

1.4.1. Teoría de ciclos, cálculo socialista y praxeología (Mises, 1912; 1920; 1949)

Entre las aportaciones de Ludwig von Mises destacan:

Teoría monetaria y de los ciclos: en su obra de 1912 estudia los orígenes y el funcionamiento del dinero. Mediante el "teorema de la regresión" resuelve la aparente circularidad en la determinación del valor del dinero, haciéndolo retroceder hasta el momento en que el bien-dinero tenía valor como mercancía ordinaria, lo que le permite extender al dinero la teoría de la utilidad marginal.

Además, sienta las bases de la teoría austriaca del ciclo, desarrollada después por Hayek, en la que se rompe la neutralidad monetaria clásica. La expansión del crédito sitúa el tipo de mercado por debajo del natural y orienta la inversión hacia procesos productivos demasiado largos (*malinvestment*), que la escasez de ahorro obliga después a liquidar.

> La teoría austriaca del ciclo (sobreinversión monetaria) se expone, junto con las demás teorías del ciclo anteriores a Keynes, en el tema A42 (§1.2.2).

Cálculo económico socialista: en 1920 Mises sostuvo que, en ausencia de propiedad privada de los medios de producción y, por tanto, de precios de mercado, el socialismo no podía realizar una asignación racional de los recursos. Esta tesis abrió el debate del cálculo socialista, en el que sus principales contradictores fueron los socialistas de mercado Oskar Lange y Abba Lerner (el planificador podía simular el mercado por ensayo y error), sobre la base teórica previa de Enrico Barone (1908).

Praxeología: en *La acción humana* (1949) sistematiza, a partir de su obra de 1933, la praxeología, es decir, el estudio de la acción humana a partir de axiomas *a priori*, como enfoque metodológico de la economía.

1.4.2. Teoría del ciclo y conocimiento disperso (Hayek, 1931; 1935; 1945)

Friedrich von Hayek, discípulo de Mises, se incorpora en 1931 a la London School of Economics y convierte el aparato austriaco en la principal alternativa teórica al keynesianismo emergente. Dentro del marco temporal de este tema destacan dos aportaciones:

Teoría austriaca del ciclo: en su obra de 1931 formaliza la intuición de Mises mediante el llamado *triángulo hayekiano*, que representa la estructura intertemporal de la producción por etapas. Además, Hayek toma de Wicksell el concepto de tipo de interés natural (Anexo A.3). De esa teoría del ciclo proceden su oposición a las políticas de estímulo de la demanda, que a su juicio solo posponen la liquidación de las inversiones erróneas, y su polémica con Keynes en 1931-1932.

Cálculo económico socialista: en 1935 reformula la tesis de Mises y, en *The Use of Knowledge in Society* (1945), ya como respuesta a Lange (1936), la convierte en un problema de conocimiento. La información relevante para la asignación está dispersa entre los agentes, es en parte tácita y solo se revela en el propio proceso de mercado a través de los precios, de modo que ningún planificador puede reunirla de antemano.

1.4.3. Otros autores de la esfera austriaca

En la Viena de entreguerras se forma además una generación que se integra en la corriente principal (Haberler, Machlup, Rosenstein-Rodan y Morgenstern, coautor con von Neumann de la teoría de juegos, A14). En el coloquio matemático de Karl Menger, hijo de Carl, se obtienen las primeras demostraciones de existencia del equilibrio general (apartado 3.4).

1.5. Conclusiones

Aspectos positivos: la revolución marginalista austriaca introdujo el valor subjetivo y el análisis marginal como núcleo de la teoría del valor, desplazando definitivamente las teorías clásicas del coste de producción y del valor-trabajo. Sus aportaciones (utilidad marginal, imputación, coste de oportunidad, papel del tiempo y la incertidumbre, competencia como proceso dinámico) resultaron fundamentales para la microeconomía posterior.

Limitaciones:

- *Metodología*: el rechazo del formalismo matemático y de la contrastación empírica, que en Menger es una preferencia por la explicación causal y en Mises se convierte en apriorismo declarado, dejó a la escuela sin los dos instrumentos que la profesión adoptó durante el periodo. Además, la negativa a determinar los precios de forma simultánea la separó de Marshall y de Walras justo cuando ambos convergían. Por estas razones, aunque sus ideas se recuperaron en los años setenta (Kirzner, Rothbard y el Premio Nobel de Hayek en 1974), la escuela ha permanecido en la heterodoxia.

- *Imputación y precios de los factores*: la teoría de la imputación de Wieser solo resuelve el problema con coeficientes de producción fijos. En cambio, cuando los factores son sustituibles, el valor del producto no puede repartirse entre ellos sin recurrir a la productividad marginal, que los austriacos no llegaron a desarrollar. Por eso, la solución general vendría de Clark, Wicksteed y Wicksell (apartado 2.5.1).

- *Teoría del capital*: el periodo medio de producción de Böhm-Bawerk, sobre el que descansa su explicación del interés, resultó ser una magnitud mal definida, porque solo es calculable bajo supuestos muy restrictivos. Además, su medición exige conocer previamente el tipo de interés que Böhm-Bawerk pretende determinar (objeción de Fisher y Knight, nunca resuelta satisfactoriamente).

2. Esfera anglosajona

La esfera anglosajona, británica en su origen, acaba fijando el canon de la microeconomía moderna. Se distingue de las demás en que el marginalismo no aparece aquí sobre un vacío teórico, sino frente a la escuela clásica más sólida y mejor establecida de Europa, la de Ricardo y Mill.

La aportación de Jevons (1871) es paralela y simultánea a la de Menger, pero con una metodología opuesta, matemática y estadística. El desarrollo de la esfera lo asume Alfred Marshall, que reintroduce el lado de la oferta que Jevons y Menger habían relegado y funda la escuela de Cambridge (Pigou, Robertson, Robinson y el propio Keynes). En Estados Unidos, una generación posterior (Clark, que se había formado en Alemania, Fisher y Knight) extiende el principio marginal a la distribución y al capital, con un tono normativo que procede de la preocupación por la reforma social.

2.1. Antecedentes: el proto-marginalismo británico

Los antecedentes británicos son numerosos pero inertes. En un medio dominado por la teoría clásica del valor-trabajo y del coste de producción, varios autores formularon con claridad la relación entre el valor y la utilidad de la última unidad sin que ello alterase el paradigma vigente.

- *Precursores británicos*: Jeremy Bentham había fundamentado el utilitarismo en el cálculo del placer y el dolor. William F. Lloyd (1833), Mountifort Longfield (1834) y Nassau Senior (1836) enunciaron el principio de utilidad decreciente y su relación con el precio, en el caso de Senior dentro de una obra que seguía siendo clásica en lo demás.

- *Bernoulli*: precursor común a las tres esferas fundadoras, recurrió ya en 1738 a una utilidad marginal decreciente para resolver la paradoja de San Petersburgo e introdujo la utilidad esperada (A10).

2.2. William Stanley Jevons (*The Theory of Political Economy*, 1871)

Formación: estudió química y matemáticas en el University College de Londres y trabajó como ensayador en la Casa de la Moneda de Sídney (1854-1859) antes de dedicarse a la economía. Fue profesor en el Owens College de Manchester y en el University College de Londres.

Importancia: Jevons, en 1871[6], parte de la utilidad marginal para reconstruir sobre ella el conjunto de la teoría del intercambio y de la distribución, con la pretensión explícita de desplazar a Ricardo y a Mill. Como Menger, aplica el análisis marginal sobre todo a la demanda.

Metodología: Jevons emplea un método matemático y estadístico, opuesto al de Menger, y sostiene que la economía, si ha de ser una ciencia, debe ser una ciencia matemática. Jevons aplica además la estadística al estudio de los precios (números índice) y del ciclo.

[6] Jevons había expuesto ya la teoría en 1862, en una comunicación a la Asociación Británica para el Progreso de la Ciencia que se publicó en 1866 y pasó inadvertida.

2.2.1. Teoría del valor: demanda, oferta y equilibrio

Al igual que Menger, Jevons defiende que el valor depende del grado final de utilidad (*final degree of utility*), es decir, de la utilidad marginal. Entiende además que existe una relación causal secuencial y unidireccional, conocida como *cadena de Jevons*, según la cual el coste de producción determina la oferta, la oferta determina el grado final de utilidad y este determina el valor. En último término, por tanto, la utilidad (la demanda) fija el precio. Esta cadena fue criticada por Marshall, Edgeworth y la escuela de Lausana, que subrayaron la determinación simultánea (y no sucesiva) de los precios mediante el equilibrio de la oferta y la demanda.

En su teoría del intercambio, Jevons llegó a la conclusión de que, si dos partes o "cuerpos comerciales" (*trading bodies*) actúan como precio-aceptantes, alcanzan un equilibrio de mercado mediante un proceso de intercambio. Su análisis se basa en:

1. *Ley de la indiferencia*[7]: en un mercado libre y abierto, una misma mercancía no puede tener en un momento dado dos precios distintos. Aplicada al consumidor, la ley implica que, a efectos de su grado final de utilidad, cualquier unidad de un bien homogéneo que consuma puede sustituirse por otra.

2. *Condición de equilibrio*: dos partes solo intercambian si ambas ganan. El intercambio continúa, por tanto, mientras las relaciones marginales de sustitución (RMS) de las dos partes, cocientes de sus utilidades marginales, difieran entre sí o de la relación de intercambio, que la ley de la indiferencia hace única. En el equilibrio, todas coinciden:

$$\mathrm{RMS}^{A}_{YX} = u'^{A}_{X}/u'^{A}_{Y} = p_X/p_Y = u'^{B}_{X}/u'^{B}_{Y} = \mathrm{RMS}^{B}_{YX}$$

2.2.2. Teoría de la producción y la distribución: el trabajo, el capital y el interés

Jevons aplica también el razonamiento marginal a la oferta de trabajo. El trabajador prolonga su jornada hasta que la utilidad del producto de la última hora iguala la penosidad (*painfulness*) de ese esfuerzo. La cantidad de trabajo queda así determinada por la igualdad entre la utilidad marginal del producto y la desutilidad marginal del trabajo.

Jevons fue además uno de los primeros en concebir el capital como tiempo. Lo entiende como el adelanto de subsistencias a los trabajadores, que supone una renuncia temporal y cuyo fruto, el interés, solo aparece más tarde. Si $F(t)$ es el producto que se obtiene al mantener invertido el capital durante un periodo t, el tipo de interés es el incremento relativo del producto al prolongar la inversión, $F'(t)/F(t)$, idea que anticipa el periodo de producción de Böhm-Bawerk (apartado 1.3.2).

2.2.3. Teoría del crecimiento (y de los ciclos): la cuestión del carbón y las manchas solares

En su libro de 1865, Jevons sostuvo que el crecimiento británico dependía de un carbón barato y agotable, y advirtió que las mejoras en la eficiencia de su uso no reducen el consumo, sino que lo aumentan, porque abaratan el servicio que presta. Esta tesis, conocida hoy como *paradoja de Jevons*, es el antecedente del efecto rebote que estudia la economía de la energía.

Jevons llevó a cabo además uno de los primeros estudios empíricos del ciclo con la **teoría de las manchas solares**, que atribuía las fluctuaciones de las cosechas y del crédito a la actividad solar (A42).

[7] Hoy se conoce como *ley del precio único*.

2.2.4. Conclusiones

Aspectos positivos: el apoyo en técnicas matemáticas y estadísticas (números índice, estudio empírico del ciclo) contribuyó a establecer estas disciplinas en la economía, y su concepción del capital como tiempo anticipa la de Böhm-Bawerk.

Limitaciones:

- *Teoría del intercambio*: F. Edgeworth demostró que el modelo de intercambio de Jevons no explicaba el precio, pues al tratarse de un intercambio entre dos partes (monopolio bilateral) el equilibrio quedaba indeterminado, indeterminación que se reduce al crecer el número de agentes.

- *Teoría de las manchas solares*: sus contemporáneos la rechazaron, y la correlación entre la actividad solar y las cosechas no se sostuvo con datos posteriores.

2.3. Sistematización: Alfred Marshall (*Principios de economía*, 1890)

Formación: estudió matemáticas en el St John's College de Cambridge (segundo *wrangler* en 1865), llegó a la economía desde la filosofía moral y fue catedrático de Economía Política en Cambridge entre 1885 y 1908.

Importancia: en *Principios de economía* (1890) desarrolló el marco analítico que constituye la base histórica de la microeconomía moderna, generalizó el análisis gráfico y se constituyó en líder de la escuela de Cambridge.

Marshall no es un continuador de Jevons, pues llegó al análisis marginal por su cuenta, a partir de Cournot y von Thünen, y reconcilia el marginalismo con la tradición clásica en lugar de sustituirla. Con Marshall, por tanto, el marginalismo se convierte en escuela neoclásica.

Metodología: Marshall se apoya en el razonamiento matemático, verbal y gráfico para postular sus teorías, si bien dio especial prioridad a las dos últimas herramientas con fines didácticos, relegando la formalización a notas y apéndices[8].

2.3.1. Teoría del valor: demanda, oferta y equilibrio

Idea: para Marshall, oferta y demanda determinan conjuntamente el precio de equilibrio, con lo que concilia la postura clásica (de oferta) y la marginalista (de demanda). Lo resume con el ejemplo de las tijeras, según el cual tan razonable sería preguntarse si corta el papel la hoja de arriba o la de abajo como preguntarse si el valor depende de la utilidad o del coste de producción.

Marshall estudia por separado los dos lados del mercado y los confronta después en el equilibrio. En la demanda le interesan la forma de la curva y el bienestar del consumidor, y en la oferta, la forma de las curvas y el papel del tiempo. La oferta forma parte de la teoría del valor porque, como la demanda, se expresa en precios del bien, ya que el precio de oferta de cada cantidad es lo que cuesta producir la última unidad. La tecnología que está detrás de esos costes y la remuneración de los factores que se deriva de ella se estudian a continuación.

Supuestos:

[8] Marshall resumió su criterio en una carta a Bowley (1906). Las matemáticas deben usarse como lenguaje abreviado y no como instrumento de investigación, traducirse después a palabras e ilustrarse con ejemplos importantes de la vida real, y solo entonces cabe "quemar las matemáticas".

▪ *Metodología*: equilibrio parcial. Marshall estudia un mercado aislado (*ceteris paribus*), con la renta y los precios de los demás bienes dados, a diferencia de Walras, que considera la interdependencia entre todos los mercados de la economía (apartado 3.2).

▪ *Agentes*:

- *Utilidad cardinal*: la utilidad es medible en unidades absolutas, de modo que los cambios en el precio o en la cantidad ofrecen una medida directa de los cambios en el bienestar.
- *Principio de marginalidad*: el valor se encuentra "en el margen" y procede de la escasez.
- *Separabilidad y aditividad*: la utilidad de un bien no depende del consumo de los demás, por lo que el análisis deja fuera la sustituibilidad y la complementariedad. Formalmente, la función de utilidad es aditivamente separable y cuasilineal, con la renta (real) no gastada en bienes, $m > 0$, como argumento lineal.
- *Utilidad marginal de la renta*: es constante, aproximación que Marshall da por buena si el bien absorbe una fracción reducida del presupuesto del consumidor.

▪ *Mercado de bienes*: competencia perfecta. Compradores y vendedores son numerosos y precio-aceptantes, y cada empresa produce con costes en forma de U.

Lado de la demanda: con estos supuestos, Marshall liga directamente la demanda de cada bien a su utilidad marginal.

▪ *Planteamiento*: el consumidor maximiza su utilidad sujeto a la restricción presupuestaria:

$$\max_{\mathbf{x},\, m} u(\mathbf{x}, m) = \sum_{i=1}^{n} u_i(x_i) + m \quad \text{s.a.:} \ \sum_{i=1}^{n} p_i x_i + m = M_0$$

donde x_i es la cantidad consumida del bien i, $u_i(x_i)$ la utilidad (separable y aditiva) que reporta, m la renta real no gastada en bienes, p_i el precio del bien i y M_0 la renta monetaria del consumidor. El lagrangiano, con el multiplicador λ asociado a la restricción presupuestaria, es $\mathcal{L} = \sum_{i=1}^{n} u_i(x_i) + m + \lambda \left(M_0 - \sum_{i=1}^{n} p_i x_i - m\right)$.

▪ *CPO*: con solución interior y la restricción agotada, se deriva el lagrangiano respecto a la cantidad de cada bien y a la renta no gastada. En el óptimo, la utilidad marginal de cada bien equivale a su precio multiplicado por la utilidad marginal de la renta, que es constante:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_i} &= \frac{du_i}{dx_i} - \lambda p_i = 0 \quad &\Rightarrow \ \frac{du_i}{dx_i} &= \lambda p_i \\ \frac{\partial \mathcal{L}}{\partial m} &= 1 - \lambda = 0 \quad &\Rightarrow \ \lambda &= 1 \end{aligned}$$

▪ *CSO*: con utilidad marginal decreciente en cada bien, $u_i''(x_i) < 0$, la función objetivo del consumidor es cóncava y la restricción es lineal, por lo que las condiciones de primer orden son suficientes para un máximo.

▪ *Solución*: de las CPO se sigue que la demanda inversa es la utilidad marginal, $p_i = du_i/dx_i$, y, al ser esta decreciente, la curva de demanda individual tiene pendiente negativa en el plano (x_i, p_i). La demanda de mercado es la suma horizontal de las individuales (para cada precio se suman las cantidades demandadas) y hereda esa pendiente negativa, que es la ley general de la demanda. Sobre este aparato construye Marshall la elasticidad-precio y la distinción entre movimientos y desplazamientos de la curva.

Lado de la oferta: la empresa, precio-aceptante, decide cuánto produce de cada bien a partir de sus costes.

- *Planteamiento*: la empresa elige la cantidad que maximiza su beneficio:

$$\max_{x_i} \pi(x_i) = p_i x_i - C(x_i)$$

donde $\pi(x_i)$ es el beneficio, p_i el precio de mercado, que la empresa toma como dado, y $C(x_i)$ el coste total de producir x_i, suma de un coste fijo y de otro variable.

- *CPO*: se deriva el beneficio respecto a la cantidad producida y se iguala a cero. La empresa amplía la producción hasta que el precio iguala el coste de la última unidad, es decir, su coste marginal:

$$\frac{d\pi}{dx_i} = p_i - \frac{dC}{dx_i} = 0 \;\Rightarrow\; p_i = \text{Cmg}(x_i)$$

- *CSO*: basta con que el coste marginal sea creciente en el óptimo, $\text{Cmg}'(x_i) > 0$, de modo que el beneficio es cóncavo en ese punto y la condición de primer orden caracteriza un máximo.

- *Solución*: la oferta de la empresa es el tramo creciente de su coste marginal, siempre que el precio cubra el coste variable medio. A corto plazo, con un número dado de empresas, la oferta de la industria es la suma horizontal de las ofertas individuales. A largo plazo, la entrada y la salida de empresas llevan el precio al mínimo del coste medio, y la forma de la oferta de la industria depende de cómo varían los costes al crecer esta.

La forma de la oferta depende, por tanto, del periodo considerado. Marshall distingue cuatro periodos, entendidos como un recurso analítico para describir las condiciones de oferta y no como periodos cronológicos:

- *Periodo de mercado*: con la oferta fija (perfectamente inelástica), el precio depende solo de la demanda.

- *Corto plazo*: al menos un factor no puede ajustarse, así que el productor puede variar la cantidad producida pero no las dimensiones de la planta. De ahí surge la distinción entre costes fijos y variables.

- *Largo plazo*: todos los factores son ajustables y el productor escoge las dimensiones de su planta. En el plano (x_i, p_i), la oferta de la industria es horizontal con costes constantes, tiene pendiente positiva con costes crecientes y pendiente negativa con costes decrecientes.

- *Periodo secular*: permite variar, además, la tecnología.

Gráficamente, los costes y la oferta de la empresa a corto plazo se representan en el plano cantidad-precio (figura A3.1):

1. *Coste variable medio*, $\text{CVme}(x)$: recoge el coste variable por unidad. Tiene forma de U por los rendimientos primero crecientes y después decrecientes del factor variable.

2. *Coste medio*, $\text{Cme}(x)$: añade el coste fijo medio al anterior, por lo que queda por encima de él y se le acerca al aumentar la producción.

3. *Coste marginal*, $\mathrm{Cmg}(x)$: recoge el coste de la última unidad producida, tiene también forma de U y corta a los dos costes medios en su mínimo.

4. *Oferta a corto plazo*: tramo del coste marginal desde el punto de *cierre*, mínimo de $\mathrm{CVme}(x)$. A precios inferiores la empresa no produce, y la oferta coincide con el eje de precios. Por encima del punto de *beneficio nulo*, mínimo de $\mathrm{Cme}(x)$, la empresa obtiene beneficios.

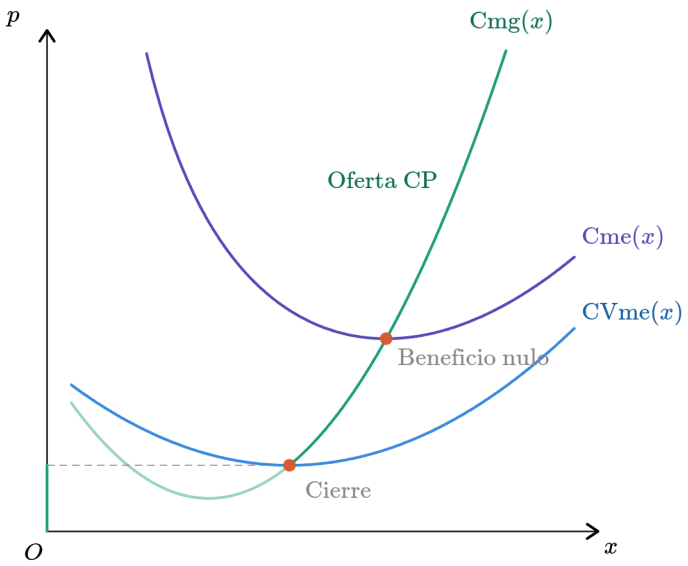

Figura A3.1: Oferta de la empresa a corto plazo. Coincide con el tramo del coste marginal situado por encima del mínimo del coste variable medio y, a precios inferiores a ese mínimo, con el eje de precios, porque la empresa no produce.

Dos fuerzas determinan, en el largo plazo, el perfil de las curvas de oferta y de costes. Las *(des)economías internas* hacen que los costes medios bajen (suban) al crecer la producción de la empresa, y las *(des)economías externas*[9] provocan ese mismo efecto al crecer la producción agregada de la industria. En general, el coste de producción pesa más en el precio cuanto más largo es el periodo analizado, mientras que la demanda pesa más cuanto más corto es.

Obsérvese que los costes decrecientes internos a la empresa (economías internas) son incompatibles con el supuesto de competencia perfecta, ya que, con un coste marginal que decrece siempre, resulta imposible maximizar el beneficio y derivar la curva de oferta de la industria. Marshall trató de salvar la dificultad atribuyendo los rendimientos crecientes a las economías externas, compatibles con la competencia, y Sraffa (1926) convertirá esa incoherencia en el punto de partida de la competencia imperfecta.

> La oferta de la empresa y de la industria en cada periodo, con la regla de cierre a corto plazo y el equilibrio a largo plazo, se desarrolla en el tema A16 (§2), y las economías y deseconomías externas, en el tema A12 (§2.4.4) y, como externalidades pecuniarias, en el tema A16 (§3.2).

[9] Marshall las asocia a la concentración de empresas en un mismo lugar, que difunde las ideas y atrae industrias auxiliares y trabajo cualificado.

Equilibrio: Marshall toma la cantidad como variable independiente y el precio como dependiente. Así, para cada cantidad, la curva de demanda señala lo máximo que pagarían los compradores y la de oferta lo mínimo que aceptarían los vendedores por llevarla al mercado. Ambas curvas se cortan en el equilibrio, la llamada *cruz marshalliana*.

Gráficamente, el equilibrio del mercado se representa en el plano cantidad-precio (figura A3.2):

1. *Curva de demanda*, $D(x)$: recoge el precio de demanda de cada cantidad, igual a la utilidad marginal, y es decreciente porque la utilidad marginal lo es.

2. *Curva de oferta*, $S(x)$: recoge el precio de oferta de cada cantidad, igual al coste marginal, y es creciente porque lo es el coste marginal.

3. *Equilibrio*: en el punto 0, con cantidad x_0 y precio p_0, el precio de demanda iguala al de oferta.

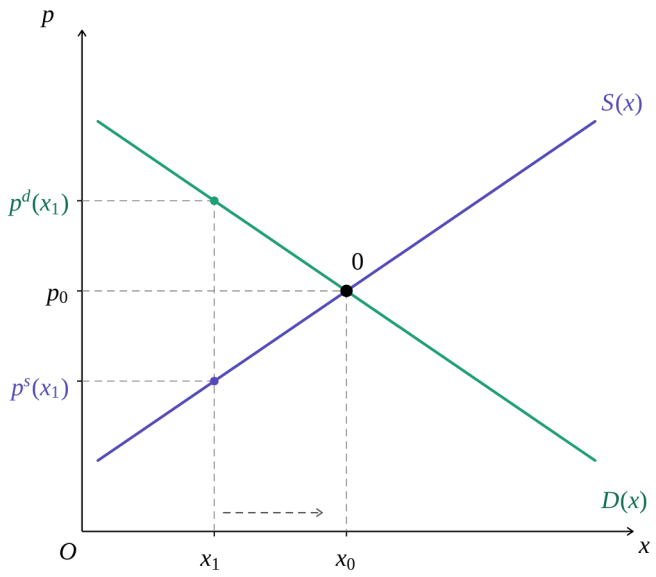

Figura A3.2: La cruz marshalliana. El precio de equilibrio resulta de la intersección de la demanda (utilidad marginal) y la oferta (coste marginal), las dos hojas de la tijera de Marshall. En x_1, el precio de demanda, $p^d(x_1)$, supera al de oferta, $p^s(x_1)$, y la cantidad aumenta hasta x_0.

Con una demanda de mercado decreciente y una oferta creciente, el cruce, si existe, es único. Marshall admitía, sin embargo, que con rendimientos crecientes la oferta a largo plazo puede ser decreciente y cortar varias veces a la demanda, de modo que se alternan equilibrios estables e inestables.

Estática comparativa: en el plano cantidad-precio (figura A3.3), el efecto de un aumento de la demanda, de $D_0(x)$ a $D_1(x)$, depende del periodo considerado. Desde el punto 0, en el periodo de mercado, con la oferta fija, $S_{\text{MP}}(x)$, solo sube el precio, hasta $p_{\text{MP},1}$. A corto plazo, la cantidad aumenta a lo largo del coste marginal, $S_{\text{CP}}(x)$, y el precio sube menos, hasta $p_{\text{CP},1}$. A largo plazo, la entrada de empresas amplía la oferta, $S_{\text{LP}}(x)$, y el precio sube aún menos, hasta $p_{\text{LP},1}$. Con costes constantes, volvería a su nivel inicial y todo el ajuste recaería en la cantidad.

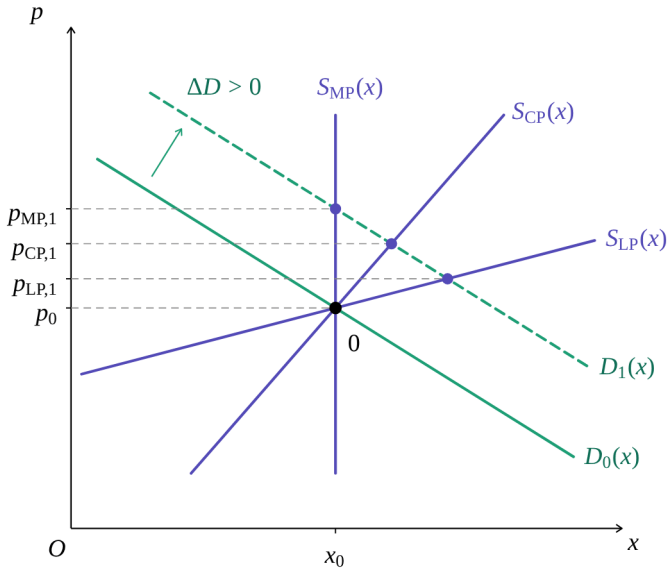

Figura A3.3: Los periodos de Marshall. Ante un aumento de la demanda, el precio se ajusta en el periodo de mercado, la cantidad, en el corto plazo, y el número de empresas, en el largo plazo.

Dinámica: el equilibrio se alcanza ajustando cantidades, no tanteando precios, porque la variable independiente es la cantidad. Partiendo de x_1, donde el precio de demanda, $p^d(x_1)$, supera al de oferta, $p^s(x_1)$, los productores amplían la cantidad ofrecida hasta x_0, punto en el que ambos precios se igualan (figura A3.2).

> El ajuste vía cantidades de Marshall y el ajuste vía precios del *tâtonnement* walrasiano, con sus condiciones de estabilidad, se desarrollan en el tema A16 (§2.4), y el tanteo en equilibrio general, en el tema A21 (§1.2.3 y §2.1.3). La cruz marshalliana y el equilibrio de la industria a corto plazo se encuentran en el tema A16 (§2.2.2).

Bienestar: del lado normativo, Marshall construye sobre la curva de demanda el *excedente del consumidor*. Dado que el precio equivale a la utilidad marginal que proporciona la última unidad consumida, las unidades intramarginales se adquieren por menos de lo que el consumidor estaría dispuesto a pagar. Esa diferencia mide en dinero la ganancia del intercambio. Como Marshall consideró que las utilidades individuales sí eran agregables, pudo sumar los excedentes individuales e interpretar el excedente de mercado como medida del bienestar colectivo.

Gráficamente, en el plano cantidad-precio (figura A3.4), el excedente del consumidor, EC, es el área entre la curva de demanda, $D(x)$, y el precio de mercado, p_0, hasta la cantidad de equilibrio, x_0.

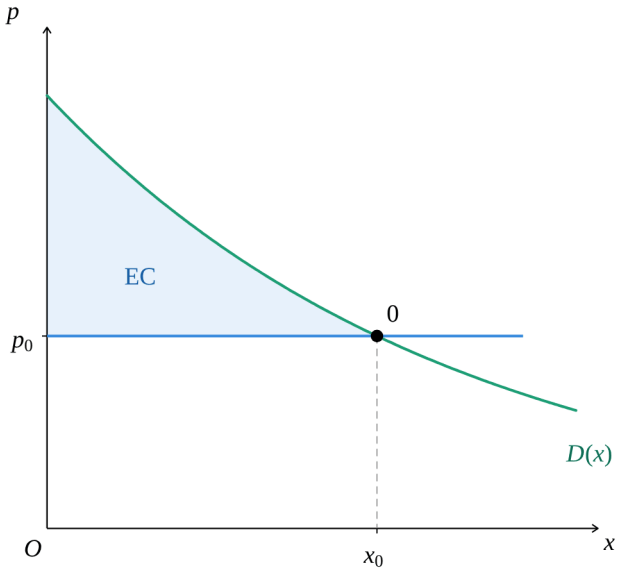

Figura A3.4: Excedente del consumidor (Marshall). El área entre la curva de demanda y el precio de mercado mide en dinero la ganancia derivada del intercambio.

La derivación del equilibrio del consumidor y de la función de demanda marshalliana, con el teorema del bien compuesto y la estática comparativa (efectos renta y sustitución, bienes normales, inferiores y Giffen, elasticidades), se desarrolla en el tema A08 (§1.3.3, §1.4 y §2), y el excedente del consumidor y su variación (VEC), junto con las medidas hicksianas de bienestar, en el tema A09 (§3.3) y, como instrumento de la economía del bienestar, en el tema A22 (§2.3).

Extensiones: por el lado de la oferta, Marshall añade la organización (que incluye la función empresarial) como cuarto factor de producción que debe ser remunerado, los conceptos de *industria* (conjunto de empresas que producen un mismo bien) y de *empresa representativa* (la que se gestiona con capacidad normal y tiene acceso normal a las economías internas y externas de su industria), y el *excedente del productor*.

Conclusiones:

▪ *Aspectos positivos*: la aportación central de Marshall es haber reconciliado la teoría clásica del coste con la marginalista de la utilidad, mostrando que la determinación del precio es simultánea.

▪ *Limitaciones*: a partir de 1900, la aportación marshalliana a la teoría de la demanda fue cuestionada por tres motivos:

- *Medición de la utilidad*: la utilidad cardinal es un supuesto muy exigente e innecesario para derivar la demanda.
- *Utilidad marginal de la renta*: suponerla constante elimina el efecto renta, de modo que el excedente del consumidor solo mide con exactitud la variación del bienestar con preferencias cuasilineales. Por la misma razón, la teoría no puede acomodar la excepción de Giffen, que el propio Marshall recogió desde la tercera edición de los *Principios* (1895).

- *Separabilidad*: la función de utilidad aditivamente separable no recoge las relaciones de complementariedad y sustituibilidad entre los bienes, que sí incorpora la función de utilidad general de Edgeworth, base de sus curvas de indiferencia.

2.3.2. Teoría de la producción y la distribución

La oferta anterior se expresa en precios del bien y toma los costes como dados. Este apartado estudia lo que está detrás de esos costes, la relación técnica entre los factores y el producto, y la remuneración de los factores que resulta de valorar ese producto a precios de mercado.

Teoría de la producción: productividad marginal: Marshall elaboró su propia versión, el producto marginal neto, es decir, el incremento del producto total ante un incremento del factor en cuestión, una vez determinadas las cantidades óptimas del resto de factores. De las curvas de productividad marginal se obtienen curvas de demanda de factores, lo que generaliza la teoría de la distribución ricardiana, que pasa a ser un caso particular de la teoría de la productividad marginal.

Teoría de la distribución: producto neto en el margen y precio de los factores: Marshall aplicó a la distribución el mismo análisis basado en el tiempo, lo que le permitió responder a las dos controversias de la teoría marginalista de la distribución:

- *Medición de la productividad marginal de cada factor*: el producto neto en el margen le permite defender que la teoría marginalista de la distribución es una teoría de la demanda de factores.

- *Determinación de los precios de los factores*: frente a la tesis marginalista de que el precio del bien final determina el de los factores, Marshall muestra que, según el marco temporal, el pago de un factor determina ese precio o depende de él. El marco temporal importa porque influye en la elasticidad de la oferta de los factores.

Con este fin Marshall acuña la *cuasirrenta*, el exceso sobre su coste de producción que obtiene un factor mientras su oferta no puede ajustarse o mientras dura una circunstancia excepcional del mercado. La cuasirrenta es una noción de corto plazo, porque con tiempo suficiente la oferta responde y el exceso se disipa. Marshall la aplica por igual a la renta de la tierra, los salarios, los beneficios y los intereses, y siempre llega al mismo resultado[10]:

- *Corto plazo*: con la oferta del factor fija, su remuneración es consecuencia del precio del bien final.

- *Largo plazo*: con una oferta muy elástica, la causalidad se invierte y la remuneración del factor determina el precio del bien final.

2.3.3. Teoría del crecimiento (y de los ciclos)

Marshall atribuye el crecimiento a la organización, que contrarresta los rendimientos decrecientes de la tierra. Según sus *Principios*, la parte que la naturaleza desempeña en la producción tiende a rendimientos decrecientes, mientras que la que desempeña el hombre tiende a rendimientos crecientes, gracias a la división del trabajo, la maquinaria y las economías externas.

[10] En la renta de la tierra matiza a Ricardo. Para la economía en su conjunto (oferta fija de tierra) la renta está determinada por el precio de los bienes, pero para una empresa o un uso concreto es un coste que entra en el precio.

En cuanto al ciclo, aceptó las ideas de Mill sobre la estabilidad de la economía (no cabe una superproducción general permanente) y atribuyó las fluctuaciones a factores psicológicos como la confianza empresarial, postura criticada por el subconsumismo de Hobson (Anexo A.2).

2.3.4. Teoría del comercio

En su ensayo de 1879, Marshall determina gráficamente la relación real de intercambio de equilibrio (véase la figura del tema B05, apartado 2.1.2). Para ello traduce la demanda recíproca de Mill a curvas de oferta neta (*offer curves*), que muestran cuántas exportaciones ofrecería un país a cambio de importaciones según cuál sea la relación real de intercambio.

> Las curvas de oferta neta y la determinación de la relación real de intercambio se derivan en el tema B05 (§2.1).

2.3.5. Teoría monetaria y bancaria: la versión de Cambridge

Marshall analizó los determinantes del nivel general de precios mediante un análisis de oferta y demanda basado en saldos de caja. Lo expuso ante la Comisión del Oro y la Plata (1887-1888) y lo difundió en sus clases, pero el enfoque solo se publicó con Pigou (1917) y con el propio Marshall (1923).

Este enfoque dio fundamentos microeconómicos a la teoría cuantitativa clásica (Hume, 1752). Su ecuación es $M = kPY$, donde k es la proporción de la renta monetaria PY que hogares y empresas desean mantener en forma de dinero. Se supone que k es estable y que la renta real está en su nivel de pleno empleo, independiente de M (por la flexibilidad del salario real y la ley de Say). En ese caso, un aumento de la cantidad de dinero produce un aumento proporcional del nivel de precios y, con él, de la renta monetaria.

> El enfoque de los saldos de caja de Cambridge se deriva en el tema A35 (§1.1.2).

2.3.6. Conclusiones

Aspectos positivos: buena parte del utillaje microeconómico vigente es marshalliano (cruz de oferta y demanda, elasticidad, excedentes, corto y largo plazo, cuasirrenta, economías externas).

Limitaciones:

- *Teoría de la empresa y los mercados*: el enfoque marshalliano de los costes y el mercado (curvas de coste en forma de U y competencia perfecta) fue ampliamente superado por la teoría de la organización industrial (TOI).

- *Teoría de la distribución*: la división en periodos es un artificio analítico sin correlato nítido en la realidad, de modo que la doctrina de las cuasirrentas no permite establecer, en cada caso concreto, si la demanda o la oferta fija la remuneración de un factor.

- *Teoría monetaria*: la versión de Cambridge descansa en la estabilidad de k y en la ley de Say, supuestos que la Gran Depresión desmintió. La *Teoría general* los sustituiría, respectivamente, por una demanda de dinero sensible al tipo de interés y por el principio de la demanda efectiva.

2.4. Evolución posterior: Edgeworth, Pigou, la escuela de Cambridge y la revolución ordinal

Los autores que siguen, sin discutir ya los fundamentos del sistema marshalliano, lo extienden o lo corrigen desde dentro. La corrección más profunda partió de la crítica de Piero Sraffa (1926) ya señalada, que condujo a la competencia imperfecta de los años treinta.

2.4.1. Teoría del intercambio puro (Edgeworth, 1881)

La teoría del intercambio puro de Francis Edgeworth, anterior a los *Principios* (1890) pero fecunda solo en desarrollos posteriores, se apoya en herramientas gráficas para desarrollar una teoría de la negociación y el intercambio. Edgeworth (1881) parte de una función de utilidad general, no separable, e introduce las curvas de indiferencia y la curva de contratos[11], que recoge una multiplicidad de asignaciones eficientes en el sentido de Pareto (véase la figura del tema A21, apartado 3.1.2).

Anticipa además, mediante el mecanismo de *recontratación*, los conceptos de coalición de agentes y de **núcleo** (bautizado así en los años cincuenta), entendido este como el conjunto de asignaciones alcanzables no bloqueadas por ninguna coalición. Su conjetura de que el núcleo se reduce a las asignaciones competitivas cuando crece el número de agentes la demostraron Debreu y Scarf (1963), como se recoge en el tema A21, apartado 3.1.3.

2.4.2. Economía del bienestar normativa: intervención y regulación (Pigou, 1920)

Arthur Pigou, alumno de Marshall y su sucesor como profesor en Cambridge, desarrolla la economía del bienestar basada en el equilibrio parcial. Sostenía que el bienestar, pese a ser un estado de ánimo subjetivo, era medible en su parte económica, la que admite medición monetaria[12].

Su trabajo principal es *The Economics of Welfare* (1920), punto de partida de la teoría de los fallos de mercado. En él generaliza las economías externas de Marshall como divergencia entre el producto neto marginal privado y el social, origen del concepto de externalidad, y propone corregirla con impuestos y subvenciones (los impuestos pigouvianos).

2.4.3. Teoría de la competencia imperfecta (Robinson, 1933; Chamberlin, 1933)

En el marco del enfoque de equilibrio parcial, destacan las obras de Joan Robinson (1933) y de Edward Chamberlin (1933) en torno a la competencia imperfecta y la competencia monopolística, respectivamente. Ambos autores llegan a resultados en parte coincidentes por caminos independientes, la primera desde la crítica de Sraffa en Cambridge y el segundo desde Harvard.

Robinson: difundió el término *monopsonio* y analizó la relación entre las magnitudes medias y marginales del ingreso y del coste, y en qué casos cabe discriminar precios.

Chamberlin: acuñó el término *competencia monopolística* y su trabajo tuvo mayor alcance que el de Robinson, porque incorporó la diferenciación del producto y los costes de venta. Formula el teorema del exceso de capacidad, que, sin embargo, descansa en una curva de costes medios neoclásica en forma de U.

[11] La llamada *caja de Edgeworth-Bowley* no aparece en Edgeworth (1881), que solo dibuja la curva de contratos. El diagrama en forma de caja lo introduce Pareto (1906) y lo populariza Bowley (1924).

[12] Su criterio es el dividendo nacional (la renta nacional), que debe ser (i) suficientemente amplio, (ii) repartido de forma más igualitaria y (iii) estable.

Estas aportaciones, todavía basadas en el paradigma neoclásico, fueron superadas por el desarrollo de la TOI, que "entró" en la "caja negra" de la empresa neoclásica. La competencia monopolística resurgió, no obstante, con el modelo de Dixit y Stiglitz (1977), base de la nueva teoría del comercio (A18, apartado 2.2).

2.4.4. Teoría ordinal de la demanda y medición del bienestar (Hicks y Allen, 1934; Hicks, 1939b)

Las objeciones a la teoría marshalliana de la demanda (medición cardinal y separabilidad) se resuelven primero en Yale, Lausana y Kiev y definitivamente en la London School of Economics.

Teoría neoclásica de la demanda del consumidor (Hicks y Allen, 1934): Fisher (1892), Pareto (1906), Slutsky (1915) y Hicks y Allen (1934) concluyen que la medición cardinal de la utilidad es innecesaria para derivar la demanda, porque basta con que los individuos ordenen las cestas. Nace así la teoría ordinal de la utilidad. Desde entonces se usan curvas de indiferencia para representar las preferencias, y su pendiente mide la relación marginal de sustitución (A08).

Cuantificación del bienestar: variación equivalente y compensatoria (Hicks, 1939b, 1941 y 1942): John Hicks emplea la variación compensatoria de la renta en su obra de 1939 y define ambas medidas de bienestar en sus artículos sobre el excedente del consumidor (1941 y 1942). Estas medidas sirvieron de apoyo a la defensa de los principios de compensación potencial (A09, A22).

2.5. Evolución posterior en Estados Unidos: Clark, Fisher y Knight

La extensión norteamericana gira en torno al problema del **agotamiento del producto**, que ocupa a Clark, Wicksteed, Flux y Wicksell entre 1888 y 1901. Además, la extensión norteamericana reformula la teoría del capital y el interés.

2.5.1. John Bates Clark (*La distribución de la riqueza*, 1899)

El autor:

- *Formación*: estudió en el Amherst College y, entre 1872 y 1875, en Zúrich y en Heidelberg, donde fue alumno del historicista Karl Knies. Desde 1895 enseñó en la Universidad de Columbia.

- *Importancia*: es el principal representante estadounidense del marginalismo y uno de los fundadores de la American Economic Association (1885), que desde 1947 concede la medalla que lleva el nombre de Clark.

- *Metodología*: separa la *estática*, el estudio de una economía competitiva en la que no cambian la población, el capital, la técnica, la organización de la industria ni los gustos, de la *dinámica*, y formula sus leyes para la primera.

Principales aportaciones:

- *Teoría del capital y el interés*: en 1888 plantea una teoría alternativa del interés. Separa el fondo abstracto de capital o *capital social*, permanente y móvil entre empleos, de los bienes de capital concretos y perecederos en que se va encarnando. Ese tránsito es la *transmigración del capital*, por la que, según Clark, la vida del capital es interminable. Los bienes concretos obtienen rentas, en cantidades absolutas, y el fondo, un interés expresado como porcentaje sobre su valor, que en estado estacionario iguala la productividad marginal del capital.

▪ *Teoría de la distribución*: en 1899 formula la teoría marginal de la distribución, según la cual, en competencia, cada factor cobra su productividad marginal.

▪ *Agotamiento del producto*: con una teoría residual como la de Ricardo, el reparto agota el producto por construcción, porque una de las rentas recoge lo que dejan las demás y los pagos a los factores suman el producto total[13]. Si, en cambio, cada factor cobra según su productividad marginal, ese resultado no está garantizado[14], y hay que probar que la suma de las remuneraciones (es decir, los costes totales) iguala al ingreso total[15]. La demostración llega en tres etapas:

1. *Clark* (1891; 1899): el primer intento de demostración es del propio Clark, pero, al razonar verbalmente, no ofrece una demostración formal. En realidad, buscaba rebatir al crítico social Henry George y defender el sistema capitalista frente al marxismo.

2. *Wicksteed y Flux* (1894): prueban que el agotamiento del producto se cumple con competencia perfecta y funciones de producción homogéneas de grado uno (rendimientos constantes a escala). Aplicando el teorema de Euler, el producto total coincide entonces con la suma de las participaciones de los factores (A11). Por ello, en un principio se creía que la existencia de rendimientos constantes a escala para cualquier nivel de producción, junto con la competencia perfecta, era necesaria para que la ganancia económica (ingresos menos costes) fuese nula y, así, se agotara el producto.

3. *Wicksell* (1900; 1901): el sueco Knut Wicksell cierra la cuestión. Al estudiar las etapas de la producción, advirtió que los rendimientos constantes a escala para cualquier nivel de producción (la homogeneidad de grado uno) son una condición suficiente, pero no necesaria. La condición necesaria y suficiente es que haya competencia perfecta y que la empresa produzca en el punto mínimo de su curva en forma de U de costes medios a largo plazo, es decir, que se den rendimientos constantes en la dimensión óptima.

Conclusiones:

▪ *Aspectos positivos*: la teoría de la productividad marginal extiende a la distribución el principio que los marginalistas habían aplicado al valor y explica con una sola regla la remuneración de todos los factores.

▪ *Limitaciones*: Clark no llegó a demostrar el agotamiento del producto, y su lectura normativa, según la cual cada factor recibe lo que crea, no se sigue de la teoría (apartado 2.6).

2.5.2. Irving Fisher (*La teoría del interés*, 1930)

El autor:

▪ *Formación*: estudió matemáticas en Yale, donde obtuvo en 1891 el primer doctorado en economía de esa universidad, bajo la dirección del físico Josiah Willard Gibbs y del sociólogo William Graham Sumner.

[13] En la teoría ricardiana, la masa salarial es el producto del salario de subsistencia por la cantidad de trabajo, y la renta es diferencial (el excedente de las tierras más fértiles sobre la tierra marginal, que no paga renta). El beneficio es lo que resta del producto total una vez deducidas ambas (A02).

[14] La teoría ricardiana es un caso particular de la teoría de la productividad marginal en la que se consideran coeficientes fijos entre capital y trabajo (la "dosis"), rendimientos decrecientes de la dosis aplicada a la tierra y la tierra como factor de producción exógeno.

[15] Si el producto se agota, queda por explicar el interés, y, si no se agota, cuándo existe un beneficio positivo. Esta cuestión motivó las teorías del capital, el interés y el beneficio.

▪ *Importancia*: Fisher recibe ideas dispersas y en parte incompatibles (la preferencia temporal de Böhm-Bawerk, la productividad del capital de Clark, la teoría cuantitativa clásica) y las integra en un sistema. Ya en su tesis (1892) había representado el equilibrio general con un modelo hidráulico.

▪ *Metodología*: combina la formalización matemática con la medición estadística (desarrolló la teoría de los números índice) y fue el primer presidente de la Econometric Society, fundada en 1930.

Principales aportaciones:

▪ *Teoría del capital y el interés*: en 1907 y 1930, Fisher toma las ideas fundamentales de Böhm-Bawerk, descarta los elementos confusos y formula la teoría del interés aceptada en la actualidad. Frente a la clasificación clásica de la renta en salarios, rentas, beneficios e intereses, sostiene desde 1906 que el valor del capital es el valor descontado de sus flujos de renta futuros. El tipo de interés, el tipo de descuento que vincula ambos, deja así de ser una porción residual de la renta para convertirse en el precio que relaciona flujos situados en distintos momentos del tiempo. Su explicación integra además dos fuerzas que sus predecesores trataban por separado, la preferencia temporal (la *impaciencia*) y la productividad del capital (las *oportunidades de inversión*), de las que salen su teoría del consumo y su teoría de la inversión.

▪ *Teoría del consumo intertemporal*: la impaciencia se formaliza en un modelo de dos periodos. El consumidor reparte su renta entre consumo presente y futuro, prestando o tomando prestado al tipo de interés, hasta que su relación marginal de sustitución entre ambos iguala $1 + r$. El consumo depende así del valor actual de toda su renta, y no solo de la del periodo. De este modelo parten las teorías de la renta normal (tema A33, apartado 2.1).

▪ *Teoría de la inversión*: las oportunidades de inversión permiten transformar consumo presente en futuro, pero el tipo de interés solo será positivo si los individuos prefieren los bienes presentes. El agente invierte mientras la *tasa de retorno sobre el coste* de un proyecto supera el tipo de interés, regla que Keynes recogerá con el nombre de eficiencia marginal del capital (tema A34, apartados 1.1 y 1.2).

▪ *Teoría monetaria y de la inflación*: Fisher añade cuatro resultados:

- *Hipótesis de expectativas puras o insesgadas* (1896): el tipo de interés a largo plazo es la media de los tipos a corto esperados (tema B24, apartado 2.3.1).

- *Ecuación de Fisher* (1896; 1930): distingue el tipo de interés nominal del real. Partiendo de $(1+i) = (1+r)(1+\pi^e)$ y despreciando el término cruzado, se aproxima como $r \approx i - \pi^e$, donde r es el tipo de interés real, i el tipo de interés nominal y π^e la tasa de inflación esperada.

- *Teoría cuantitativa* (1911): la versión transacciones se debe a Fisher. Esta versión reinterpreta en sentido causal la identidad contable del gasto (ecuación de cambio, formulada por Newcomb en 1885), $MV = PT$, donde V y T se consideran independientes de M (salvo en los periodos de transición). Por ello, la relación entre cantidad de dinero y precios es directa y proporcional y el dinero resulta neutral (tema A35, apartado 1.1.1).

- *Deflación por deudas* (1933): rompe con esa neutralidad. Ante una situación de sobreendeudamiento, la liquidación masiva de activos contrae precios y crédito. Como las deudas están fijadas en términos nominales, la deflación eleva su valor real, de manera que el esfuerzo colectivo por desapalancarse aumenta la carga real de la deuda en lugar de reducirla.

Conclusiones:

▪ *Aspectos positivos*: su formulación del interés es la que ha llegado a la teoría actual y microfundamenta el mercado de fondos prestables y los modelos intertemporales de consumo e inversión. De ella procede además el *teorema de separación de Fisher*, según el cual, con mercados de capitales perfectos, la empresa maximiza su valor presente con independencia de las preferencias temporales de sus propietarios, base de la teoría financiera de la inversión.

▪ *Limitaciones*: la hipótesis de expectativas puras es rechazada de forma sistemática por los datos, que documentan primas por plazo no nulas y variables. La estabilidad de V y de T, por su parte, es un supuesto exigente, que la propia Depresión desmintió y que el mismo Fisher hubo de matizar en 1933.

> El modelo de elección intertemporal de Fisher (consumo y ahorro con tipo de interés) y el teorema de separación se desarrollan en el tema A29 (§1.4 y §2.4.2).

2.5.3. Frank Knight (*Riesgo, incertidumbre y beneficio*, 1921)

El autor:

▪ *Formación*: se doctoró en Cornell en 1916, con la tesis de la que procede su libro de 1921. Enseñó en la Universidad de Iowa (1919-1927) y, desde 1927, en la de Chicago.

▪ *Importancia*: es, con Jacob Viner, una de las figuras fundadoras de la escuela de Chicago, y entre sus alumnos figuran Milton Friedman, George Stigler y James Buchanan.

▪ *Metodología*: hace de la teoría de los precios el núcleo de la disciplina, pero advierte de los límites del método científico en economía, frente al empirismo de los institucionalistas.

Principales aportaciones:

▪ *Riesgo e incertidumbre*: distingue el *riesgo*, cuyas probabilidades son medibles y que, por tanto, puede asegurarse, de la *incertidumbre*, que no admite medición ni seguro (tema A10).

▪ *Teoría del beneficio*: sostiene que los beneficios existen porque hay incertidumbre. El beneficio remunera, por tanto, a quien asume la incertidumbre genuina, mientras que el riesgo, que el seguro convierte en un coste más, no lo explica.

Conclusiones:

▪ *Aspectos positivos*: la teoría da al beneficio un fundamento propio, distinto del interés del capital y del salario de dirección, y sitúa al empresario en el centro de la explicación.

▪ *Limitaciones*: la frontera entre riesgo e incertidumbre es difícil de trazar. Con probabilidades subjetivas, toda incertidumbre puede tratarse como riesgo, y la distinción solo recupera contenido con la evidencia de aversión a la ambigüedad (tema A10).

2.6. Conclusiones

Aspectos positivos: la esfera anglosajona fija el canon instrumental de la microeconomía, y el recurso a herramientas gráficas se ha mantenido en la disciplina hasta hoy. Su rama norteamericana aporta, además, la teoría de la distribución basada en la productividad marginal, con la que el sistema neoclásico adquiere el carácter cerrado del que carecía, y la teoría del interés de Fisher.

Limitaciones:

- *Gran Depresión*: provocó el colapso de muchas de las ideas clásicas subyacentes a estas teorías (señaladamente la ley de Say y el ajuste automático al pleno empleo). El colapso obligó a introducir elementos teóricos adicionales, como la demanda efectiva de Keynes y, más tarde, las rigideces o la información imperfecta.

- *Equilibrio parcial*: el enfoque presenta importantes limitaciones, que hicieron que los modelos de equilibrio general, basados en las aportaciones de Léon Walras, lo superasen progresivamente.

- *Productividad marginal*: la teoría arrastra desde el principio la objeción de la medición, dado que el producto resulta de un esfuerzo conjunto de los factores. Además, su uso como justificación normativa de la distribución existente, explícito en Clark, excede lo que el aparato analítico permite sostener.

3. Esfera franco-italiana

La tercera esfera es la que menos se parece a las dos anteriores, puesto que nace en la tradición francesa de las matemáticas y la ingeniería aplicadas a los problemas económicos (Cournot y Dupuit), y no en las facultades de economía. Su rasgo distintivo es el uso sistemático de las matemáticas.

Sobre esa base, Léon Walras, desde la cátedra de Lausana, define la utilidad marginal como *rareté* e incorpora el lado de la oferta. Ante todo, sustituye la determinación causal y unidireccional del precio por un sistema de ecuaciones simultáneas que recoge la interdependencia de todos los mercados (la teoría del equilibrio general). La continuidad en Lausana la asegura su sucesor, Vilfredo Pareto.

3.1. Antecedentes: el proto-marginalismo francés

Los precursores franceses, matemáticos e ingenieros de formación, llegan a la economía desde problemas prácticos de fijación de precios y de evaluación de obras públicas. Ninguno de los dos formuló una teoría completa del valor.

3.1.1. Antoine Augustin Cournot (1838)

En su obra de 1838, Cournot no presentó una teoría de la utilidad marginal. Fue, sin embargo, el primero en expresar la demanda como una función decreciente del precio (*loi du débit*) y en emplear instrumentos marginales para analizar la empresa. Resolvió así el problema del monopolista, que iguala ingreso marginal y coste marginal, y el del duopolio con producto homogéneo, cuyo equilibrio lleva hoy su nombre (A19, apartado 1.1).

3.1.2. Jules Dupuit (1844)

La relación inversa entre precio y cantidad demandada aparece también en Jules Dupuit, que además la vincula con la utilidad. Sostiene que la curva de demanda es la propia curva de utilidad marginal, expresada en dinero, y que su pendiente es negativa porque la utilidad marginal decrece. De ahí deriva la medición del bienestar del consumidor mediante el excedente (su *utilité relative*), que Marshall retomará (apartado 2.3.1), y la tarificación discriminatoria de puentes y ferrocarriles (1849).

3.2. Léon Walras (*Elementos de economía política pura*, 1874)

Formación: estudió ingeniería en la Escuela de Minas de París, sin concluirla, y fue periodista, empleado de ferrocarril y gestor bancario antes de obtener en 1870 la cátedra de Economía Política de Lausana.

Importancia: de sus ideas nació la teoría del equilibrio general, recogida en *Elementos de economía política pura* (1874), que integra no solo el papel de la demanda (utilidad marginal), sino también el de la oferta (coeficientes de producción y, desde la tercera edición, productividad marginal).

Cournot y Dupuit manejaban el cálculo diferencial, pero resolvían problemas parciales sin salir del mercado que estudiaban. Walras, en cambio, convierte en objeto del análisis la interdependencia de todos los mercados, resuelta simultáneamente mediante un sistema de ecuaciones[16], y pasar del problema al sistema lo hace fundador.

Metodología: Walras hace un uso intensivo de las matemáticas, como antes Cournot, y encabeza la escuela de Lausana, de la que forma parte también Pareto.

3.2.1. Teoría del valor: demanda, oferta y equilibrio

Idea: Walras llama *rareté*[17] a la utilidad marginal y, a diferencia de Jevons y Menger, en quienes esta establece una relación causal unidireccional con el valor, la inserta en un sistema de interrelaciones entre todos los mercados.

Como Marshall, Walras plantea un programa para cada lado del mercado, pero los resuelve a la vez para todos los mercados. La oferta pertenece aquí a la teoría del valor por la misma razón que en Marshall, porque se expresa en precios. La tecnología y el papel del empresario que están detrás de ella se estudian a continuación.

Supuestos:

- *Metodología*: equilibrio general. A diferencia de Marshall, que estudia un mercado aislado (apartado 2.3.1), Walras considera la interdependencia entre todos los mercados de la economía y los resuelve a la vez.

- *Agentes*: en la aportación inicial de Walras, los consumidores tienen una función de utilidad cardinal y aditiva, al igual que en Marshall, y son propietarios de los servicios productivos, que ofrecen en función de los precios.

- *Tecnología*: en el modelo con producción, cada bien se obtiene con *coeficientes de producción* fijos, de modo que una unidad del bien i requiere a_{ji} unidades del servicio productivo j (de la tierra, del trabajo o del capital).

- *Mercados*: competencia perfecta. Consumidores y empresas toman como dados los precios de los productos y de los servicios.

Lado de la demanda: el modelo parte del intercambio puro, en el que cada uno de los $h = 1, 2, \dots, H$ consumidores decide su cesta de consumo.

- *Planteamiento*: cada consumidor maximiza su utilidad, sujeto a su restricción presupuestaria:

$$\max_{\mathbf{x}^h} u^h(\mathbf{x}^h) = \sum_{i=1}^{n} u_i^h(x_i^h) \quad \text{s.a.:} \sum_{i=1}^{n} p_i x_i^h = M_0^h$$

donde $\mathbf{x}^h$ es la cesta de consumo del consumidor h, x_i^h la cantidad del bien i en dicha cesta, $u_i^h(x_i^h)$ la utilidad (cardinal) que le reporta, cuya derivada es la *rareté* del bien, p_i el precio del bien i y M_0^h su renta. El lagrangiano, con el multiplicador λ^h asociado a la restricción presupuestaria, es $\mathcal{L}^h = \sum_{i=1}^{n} u_i^h(x_i^h) + \lambda^h \left(M_0^h - \sum_{i=1}^{n} p_i x_i^h\right)$.

[16] Habían advertido antes la interdependencia entre sectores, sin expresarla en un sistema de ecuaciones, Quesnay (*Tableau économique*), Cournot, von Thünen (interdependencia espacial) y Smith (interrelación de los procesos de mercado).

[17] El término procede de su padre, Auguste Walras, que en *De la naturaleza de la riqueza y del origen del valor* (1831) explicaba el valor por la escasez.

▪ *CPO*: con solución interior y la restricción agotada, se deriva el lagrangiano respecto a la cantidad de cada bien y se elimina el multiplicador. Para cada uno de los consumidores, la tasa de valoración subjetiva de cada par de bienes (la relación marginal de sustitución, cociente de sus *raretés*) debe coincidir con la tasa de valoración objetiva que ofrece el mercado (el precio relativo):

$$\frac{\partial \mathcal{L}^h}{\partial x_i^h} = \frac{du_i^h}{dx_i^h} - \lambda^h p_i = 0 \;\Rightarrow\; \text{RMS}_{ji}^h = \frac{du_i^h/dx_i^h}{du_j^h/dx_j^h} = \frac{p_i}{p_j} \quad \text{para todo par de bienes } i,\, j$$

▪ *CSO*: con utilidad aditiva y *rareté* decreciente en la cantidad consumida de cada bien, la función de utilidad es cóncava y la condición de tangencia caracteriza un máximo.

▪ *Solución*: las CPO y la restricción presupuestaria dan las demandas individuales $x_i^h(\mathbf{p}, M_0^h)$, que dependen de todos los precios. Su suma es la demanda de mercado de cada bien, de modo que la interdependencia entre mercados está presente desde el propio comportamiento del consumidor.

Lado de la oferta: el modelo se amplía después a la producción, en la que cada empresa toma como dados los precios de los productos y de los servicios productivos.

▪ *Planteamiento*: la empresa maximiza su beneficio:

$$\max_{x_i} \pi_i = \left(p_i - \sum_{j=1}^{m} a_{ji}\, p_j^z \right) x_i$$

donde x_i es la producción del bien i, a_{ji} el coeficiente de producción del servicio j en el bien i y p_j^z el precio de ese servicio, de modo que $\sum_j a_{ji}\, p_j^z$ es el coste unitario de producción.

▪ *CPO*: el beneficio es lineal en la producción, de manera que su derivada no depende de la cantidad producida. Por ello, solo cabe una producción positiva y finita si el precio iguala el coste unitario, y el empresario no obtiene así ni beneficio ni pérdida:

$$\frac{d\pi_i}{dx_i} = p_i - \sum_{j=1}^{m} a_{ji}\, p_j^z = 0 \;\Rightarrow\; p_i = \sum_{j=1}^{m} a_{ji}\, p_j^z$$

▪ *CSO*: con precio igual al coste unitario, cualquier escala de producción es óptima y la producción de cada bien queda determinada por la demanda.

▪ *Solución*: la oferta de cada bien es horizontal al nivel de su coste unitario, y la demanda de cada servicio productivo, $\sum_i a_{ji}\, x_i$, se deriva de la producción de todos los bienes.

Equilibrio: el equilibrio general es un vector de precios que vacía a la vez todos los mercados. Si $Z_i(\mathbf{p})$ es el exceso de demanda del bien i, la diferencia entre su demanda y su oferta de mercado a los precios $\mathbf{p}$, el equilibrio exige $Z_i(\mathbf{p}^*) = 0$ para todo i. En el modelo con producción, el sistema reúne cuatro grupos de ecuaciones: las demandas de los productos, la igualdad entre el precio y el coste unitario de cada uno, las ofertas de servicios productivos y la igualdad entre cada oferta y su demanda derivada. Walras amplía después el modelo a la capitalización y al dinero. Su sistema permite enunciar cuatro resultados, de desigual solidez:

▪ *Ley de Walras*: sean cuales sean los precios, los excesos de demanda valorados a esos precios suman cero en el conjunto de mercados, $\sum_i p_i Z_i(\mathbf{p}) = 0$.

- *Existencia*: con un bien como numerario y descontando la ecuación que la ley de Walras hace redundante, el número de ecuaciones e incógnitas coincide. El recuento sugiere la existencia de un vector de precios que vacía simultáneamente todos los mercados, pero no la demuestra.

- *Unicidad*: Walras no la garantiza, pues pueden existir varios vectores de precios de equilibrio en ausencia de condiciones como la sustituibilidad bruta.

- *Estabilidad*: Walras sostiene que el tanteo de precios (*tâtonnement*) converge, con un argumento heurístico según el cual el efecto directo de cada revisión de precio sobre su propio mercado domina a los efectos indirectos sobre los demás. Sin embargo, la convergencia solo se demostró después bajo condiciones como la sustituibilidad bruta, y Scarf (1960) mostró que en general no está garantizada.

Gráficamente, en el plano precio-exceso de demanda del bien i, con los demás precios dados (figura A3.5), $Z_i(p_i)$ es decreciente y se anula en p_i^*. A un precio mayor hay exceso de oferta y el tanteo lo baja, y a uno menor, exceso de demanda y lo sube, hasta p_i^*.

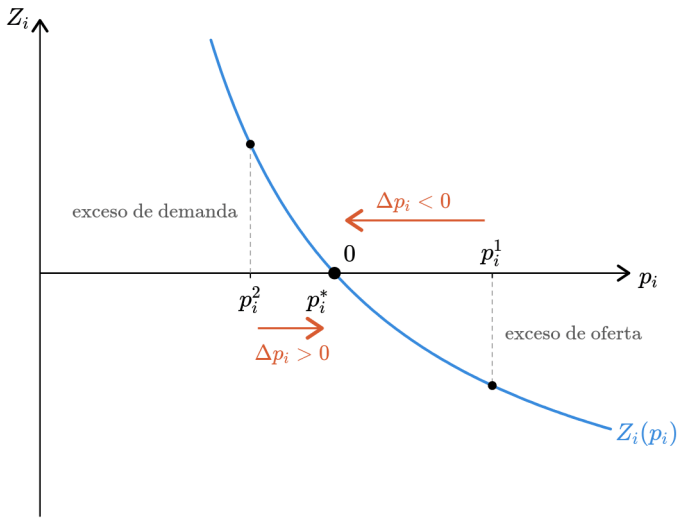

Figura A3.5: Equilibrio de Walras con la función de exceso de demanda. Con los demás precios dados, el exceso de demanda del bien i, $Z_i(p_i)$, es decreciente y se anula en el precio de equilibrio, p_i^*. A un precio mayor, p_i^1, hay exceso de oferta y el tanteo lo baja, y a uno menor, p_i^2, exceso de demanda y lo sube.

Dinámica: Walras estudia además el modo en que se alcanza el equilibrio, el *proceso de tâtonnement*[18]. Se anuncian unos precios para todos los mercados. Si, a esos precios, hay exceso de oferta o de demanda en alguno de los mercados, no se realiza ninguna transacción y los precios se revisan. Suben en los mercados con exceso de demanda y bajan en los de exceso de oferta, $\Delta p_i = \lambda_i Z_i(\mathbf{p})$, donde $\lambda_i > 0$ es la velocidad de ajuste del precio, hasta que todos se vacían (figura A3.5). Este tanteo por prueba y error, con información completa y sin costes de transacción, hace del precio la variable de ajuste, mientras que en Marshall lo es la cantidad. La tradición posterior lo personificó en un agente ficticio, el *subastador walrasiano*.

[18] Walras describe el tanteo a partir del funcionamiento de las bolsas y, desde la cuarta edición (1900), con "bonos" (*bons*) que impiden producir fuera del equilibrio (Walker, 1987).

Extensiones: el planteamiento en términos de demandas netas, con consumidores que parten de una dotación inicial de bienes cuyo valor es su renta M_0^h, procede ya del propio Walras.

> La derivación del modelo walrasiano (consumidores, empresas, ley de Walras, numerario y *tâtonnement*) se desarrolla en el tema A21 (§1.2), y la existencia (teorema del punto fijo de Brouwer), la unicidad y la estabilidad del equilibrio general, en el mismo tema (§2.1).

Conclusiones:

▪ *Aspectos positivos*: Walras convierte la determinación de los precios en un problema matemático bien planteado, un sistema de ecuaciones simultáneas con su recuento de ecuaciones e incógnitas y la ley de Walras, que la teoría del equilibrio general conserva.

▪ *Limitaciones*: el recuento de ecuaciones no asegura precios y cantidades no negativos, y la existencia y la estabilidad del equilibrio quedan como problemas abiertos para la teoría posterior.

3.2.2. Teoría de la producción y la distribución

El lado de la oferta ya expuesto se expresa en precios y toma como dados los coeficientes de producción. Este apartado estudia lo que está detrás, el papel del empresario y la tecnología, y la remuneración de los servicios productivos:

Agente-productor: Walras no trata la actividad empresarial como un factor aparte con remuneración propia, al modo de Marshall, sino como un papel que cualquier agente-consumidor puede desempeñar. En equilibrio, el empresario walrasiano no obtiene beneficio ni pérdida (*ni bénéfice ni perte*), porque los precios de los productos igualan sus costes de producción.

Coeficientes de producción: en sus primeras ediciones, Walras considera coeficientes de producción fijos, es decir, una tecnología sin sustituibilidad entre factores, lo cual presentaba evidentes limitaciones. Por ello, a partir de la tercera edición (1896) y, sobre todo, en la cuarta (1900), incorpora coeficientes variables y los fundamentos de la teoría de la productividad marginal[19].

Distribución: la remuneración de los factores no requiere una teoría aparte. Los precios de los servicios productivos (la renta de la tierra, el salario y el interés) se determinan en el mismo sistema de ecuaciones, a la vez que los de los productos.

3.2.3. Implicaciones de política económica

Walras veía su economía pura como instrumento de política económica y defendía que el Estado crease por ley mercados perfectamente competitivos, por lo que no era un defensor convencido del *laissez faire*. En su economía social propuso además la nacionalización de la tierra, cuya renta financiaría al Estado y permitiría suprimir los impuestos (1896).

[19] Con coeficientes variables la frontera de posibilidades de producción es suave y, si las intensidades factoriales difieren entre sectores, estrictamente cóncava (con coeficientes fijos es cóncava, pero quebrada).

3.2.4. Conclusiones

Aspectos positivos:

▪ *Interdependencia*: la teoría del equilibrio general supera el enfoque de equilibrio parcial de Marshall.

▪ *Economía del bienestar*: la teoría del equilibrio general sentó las bases del desarrollo teórico de la economía del bienestar.

▪ *Vigencia*: el enfoque de Lausana resulta especialmente fértil en macroeconomía (modelos de equilibrio general aplicado o computable y modelos dinámicos estocásticos de equilibrio general, DSGE).

Limitaciones: el proceso de ajuste, el *tâtonnement*, exige supuestos muy restrictivos (centralización de las ofertas y demandas, información completa, ausencia de costes de transacción y de intercambios fuera del equilibrio).

3.3. Sistematización: el sistema paretiano (Pareto, 1906)

Formación: Vilfredo Pareto se graduó como ingeniero en Turín (1870), dirigió empresas ferroviarias y siderúrgicas italianas y sucedió a Walras en la cátedra de Lausana en 1893.

Pareto depura el sistema de Walras del supuesto de utilidad cardinal que lo lastraba y le añade un criterio de evaluación normativa de las asignaciones. Con ello, el equilibrio general se convierte en el fundamento de la economía del bienestar. Pareto hace aportaciones de tres órdenes:

1. *Utilidad ordinal*: sustituye la utilidad cardinal por la ordinal (apartado 2.4.4), porque el bienestar individual no es medible en unidades absolutas ni comparable entre personas. Ya en su obra de 1896 había acuñado el término *ofelimidad* para designar la aptitud de un bien para satisfacer un deseo, legítimo o no, y distinguirla del sentido corriente de "utilidad".

2. *Criterios de Pareto*: los formula en su obra de 1906, en su intento de desarrollar una teoría del bienestar con un mínimo de juicios de valor (sin comparaciones interpersonales de utilidad). Según estos criterios, una asignación es eficiente en el sentido de Pareto cuando no es posible mejorar la situación de un agente sin empeorar la de algún otro. Se desarrollan en A21 y A22.

3. *Distribución de la renta*: en el terreno empírico, formula la ley que también lleva su nombre (1896). Según esta ley, el número de perceptores con una renta superior a un nivel dado disminuye de forma potencial con ese nivel, con un exponente parecido en distintos países y épocas.

3.4. Evolución posterior: difusión y formalización

A diferencia de las demás esferas, la evolución posterior de la escuela de Lausana no aporta nuevas tesis. Consiste en difundir el sistema de Walras y Pareto y en darle el rigor formal que le faltaba, en cinco aportaciones:

1. *Extensión a la economía planificada* (Barone, 1908): Enrico Barone, seguidor de Pareto, aplica las ecuaciones del equilibrio general a una economía colectivista. Muestra que un ministerio de la producción que quisiera maximizar el bienestar tendría que resolver el mismo sistema que el mercado competitivo, con precios iguales al coste mínimo de producción. Su trabajo es el punto de partida formal del debate sobre el cálculo económico en el socialismo (apartado 1.4.1).

2. *Simplificación y difusión* (Cassel, 1918b): la obra de Walras, escrita en francés y en un lenguaje matemático poco frecuentado, resultaba inaccesible. Gustav Cassel la rescata y simplifica su presentación, prescindiendo de la utilidad y partiendo directamente de las funciones de demanda. De ahí que el modelo se conozca a veces como modelo Walras-Cassel.

3. *Existencia del equilibrio* (Schlesinger y Wald, 1933-1936): contar ecuaciones e incógnitas no garantiza que el sistema tenga una solución con precios y cantidades no negativos. En el coloquio matemático de Viena, Karl Schlesinger reformula el sistema con desigualdades, de modo que un factor sobrante tenga precio nulo, y Abraham Wald ofrece la primera demostración rigurosa de existencia.

4. *Demostración general* (Arrow y Debreu, 1954; McKenzie, 1954): ya fuera del periodo, la existencia se demuestra de forma general. El desarrollo del equilibrio general corresponde al tema A21.

5. *Formalización de la demanda y del bienestar*: en la estela del método de Lausana, Ragnar Frisch axiomatiza la teoría de la demanda (1926; 1932), y Abram Bergson (1938) y Paul Samuelson (1947) responden con la función de bienestar social a la indeterminación de los criterios paretianos (A22). La formalización llega también al análisis dual, con el lema de Hotelling (1932), la identidad de Roy (1942) y el lema de Shephard (1953) (A09).

3.5. Conclusiones

La esfera franco-italiana da el paso del equilibrio parcial al general. Cournot había expresado la demanda como función decreciente del precio y Dupuit la había ligado a la utilidad marginal, pero ambos resolvían problemas de un solo mercado. Walras, en cambio, convierte la interdependencia de todos los mercados en el objeto del análisis y determina los precios con un sistema de ecuaciones simultáneas, que se equilibra por tanteo de precios y no, como en Marshall, de cantidades. Su recuento de ecuaciones e incógnitas sugiere la existencia del equilibrio sin demostrarla.

Pareto libera el sistema de la utilidad cardinal y le añade un criterio de eficiencia sin comparaciones interpersonales, con lo que el equilibrio general sirve también para juzgar asignaciones. No obstante, por su lenguaje y por la falta de una demostración de existencia, la obra de Lausana tardó en difundirse, y la microeconomía de entreguerras se construyó sobre el aparato marshalliano.

Conclusiones

La escuela neoclásica consolida la ciencia económica que habían fundado los clásicos. Somete el valor y la distribución a un mismo principio, el marginal, y crea los instrumentos que la disciplina todavía emplea. El desenlace del periodo (1871-1936) es, sin embargo, desigual según la esfera. La esfera austriaca acaba fuera de la ortodoxia por razones metodológicas, aunque lega el subjetivismo del valor y el coste de oportunidad. La anglosajona proporciona el aparato instrumental y, con su rama norteamericana, cierra el sistema por el lado de la distribución y del capital, y la franco-italiana aporta el armazón lógico del equilibrio general, que tardará medio siglo en imponerse. Desde los años treinta, lo que la corriente principal toma de las tres esferas forma ya un cuerpo teórico unificado, contra el que se dirigirá la crítica de Keynes.

Los instrumentos nacidos en este periodo sostienen buena parte del análisis actual. El principio marginal, aplicado al consumidor, da la demanda marshalliana y, con la utilidad ordinal de Hicks y Allen, la teoría moderna de la demanda (temas A08 y A09). Aplicado a la empresa, da la oferta competitiva de Marshall y, con la diferenciación del producto, la competencia monopolística de Chamberlin (temas A16 y A18). La interdependencia de todos los mercados que planteó Walras y el criterio de eficiencia de Pareto se convierten después en las teorías del equilibrio general y del bienestar (temas A21 y A22). En macroeconomía, la teoría cuantitativa de Fisher y de Cambridge es el antecedente de la demanda de dinero (tema A35), y la teoría austriaca, el de las explicaciones monetarias del ciclo (tema A42). Por último, la Gran Depresión, que desmiente el ajuste automático al pleno empleo, abre el periodo keynesiano (tema A04).

El tipo de interés natural de Wicksell ha vuelto a la política monetaria como tipo neutral no observable, cuya caída en las principales economías avanzadas durante las últimas décadas estiman Holston, Laubach y Williams (2017). Asimismo, la deflación por deudas de Fisher inspira los modelos de desapalancamiento con trampa de liquidez que siguieron a la crisis financiera de 2008 (Eggertsson y Krugman, 2012). La ley de Pareto, por su parte, sirve de base para estimar la concentración de las rentas más altas a partir de datos fiscales (Atkinson, Piketty y Saez, 2011).

Anexos

A.1. Diferencias entre la escuela clásica y la neoclásica

Aunque la continuidad entre ambas escuelas es notable (de ahí el término "neoclásico"), conviene fijar los puntos en que el marginalismo rompe con el esquema clásico, casi todos ellos relativos a la teoría de la distribución:

1. *Capital como factor*: en materia de distribución, clásicos y marginalistas difieren notablemente. Para los clásicos el capital no es un factor originario, sino trabajo acumulado (un medio de producción fabricado con trabajo anterior), y su remuneración, el beneficio, se explica como un excedente sobre los salarios y no por una contribución productiva propia. Para los marginalistas, en cambio (salvo Böhm-Bawerk, que explica el interés por el agio), el capital es un factor productivo más y recibe una remuneración determinada, como la de los demás factores, por su productividad marginal.

2. *Capitalista y empresario*: frente a los clásicos, los marginalistas separan la figura del capitalista de la del empresario, de modo que el pago al primero (tipo de interés) es distinto del que corresponde al segundo (beneficios). Así, en Marshall el empresario (la organización) pasa a concebirse como un cuarto factor de producción.

3. *Remuneración de los factores y explotación*: cada factor percibe una retribución acorde con su productividad marginal (incluso por debajo de la subsistencia), lo que, para sus defensores, daba respaldo al sistema capitalista y descartaba la noción marxista de explotación (apartado 1.3.2).

En la teoría del valor, Jevons y Menger sustituyeron el énfasis clásico en la oferta por un énfasis igualmente unilateral en la demanda, que Marshall y Walras corrigieron con la determinación simultánea (apartados 2.3.1 y 3.2.1).

A.2. Críticas al núcleo teórico: institucionalismo americano y keynesianismo

A diferencia del historicismo alemán, que discutía el método (tema A01, apartado 2.2.2), las corrientes de este anexo discuten el contenido, es decir, el supuesto de racionalidad individual, la estabilidad de las instituciones y, en el caso de Keynes, el corolario macroeconómico del sistema.

A.2.1. Institucionalismo americano: Veblen, Commons y Mitchell

Thorstein Veblen inicia la rama heterodoxa estadounidense que se suele denominar **institucionalismo**. Su crítica es doble:

- *Plano metodológico*: reprocha a la teoría ortodoxa su carácter estático y su concepción del individuo como maximizador aislado, y reclama una economía evolutiva, apoyada en la antropología y la psicología, que estudie el cambio de las instituciones.

- *Plano sustantivo*: en la *Teoría de la clase ociosa* (1899) muestra que buena parte del consumo es ostentoso (*conspicuous consumption*) y depende de la comparación social, idea que sobrevive en el análisis microeconómico actual bajo la denominación de bienes Veblen (véase el tema A08).

Veblen acuñó además el término "neoclásico" (1900) para poner énfasis en la ascendencia clásica de la economía marshalliana, a pesar de que era contrario a ambos paradigmas. Junto a él, otros institucionalistas ponen énfasis en aspectos no tratados por los llamados neoclásicos. John R. Commons estudia la transacción como unidad elemental de análisis y el papel del derecho y de la negociación colectiva, y anticipa así la economía de los costes de transacción (A15). Wesley C. Mitchell, por su parte, cofunda en 1920 el National Bureau of Economic Research (NBER) e impulsa el estudio empírico y estadístico del ciclo económico (A42).

A.2.2. Pensamiento keynesiano

El crítico decisivo del pensamiento neoclásico fue John Maynard Keynes, discípulo de Marshall en Cambridge. Keynes amplió el término "clásicos", acuñado por Marx para Ricardo y sus predecesores, para englobar en él también a sus seguidores (Mill, Marshall, Edgeworth, Pigou), es decir, prácticamente todas las aportaciones anteriores a la *Teoría general* (1936). Su crítica no se dirige a la microeconomía marginalista, que en buena medida acepta, sino a su corolario macroeconómico, la vigencia de la ley de Say y la tendencia automática al pleno empleo. La *Teoría general* cierra el periodo examinado en este tema y se estudia en el tema A04.

A.2.3. Institucionalismo en sentido amplio: Galbraith, Myrdal y Schumpeter

Tres autores que desbordan el marco temporal del tema suelen agruparse como *cuasiinstitucionalistas* (Landreth y Colander, 2002):

- *John Kenneth Galbraith*: centra el análisis en el poder y las instituciones (poder compensatorio, tecnoestructura) antes que en el mecanismo de precios.

- *Gunnar Myrdal*: miembro de la escuela de Estocolmo (Anexo A.3), critica los juicios de valor implícitos en la teoría y formula el principio de causación circular acumulativa (B28).

- *Joseph Schumpeter*: formado con Böhm-Bawerk y admirador de Walras, objeta que el análisis de equilibrio no puede dar cuenta del cambio. Schumpeter sitúa además al empresario innovador y la destrucción creativa en el centro del crecimiento (A42, A44).

A.2.4. Otros críticos del periodo

El tema deja planteadas además tres críticas puntuales al pensamiento neoclásico:

- *Henry George* (1879): partidario de un impuesto único sobre la renta de la tierra, fue el blanco de la demostración de Clark del agotamiento del producto (apartado 2.5.1).

- *John A. Hobson*: principal representante del subconsumismo británico, cuya tesis rechazó Marshall, influyó en Keynes (A42).

- *Rudolf Hilferding* (1904): su réplica marxista sostiene que la teoría marginalista, al partir de valoraciones subjetivas, renuncia a explicar las relaciones sociales de producción (apartado 1.3.2).

A.3. La escuela de Estocolmo (Wicksell, 1898; Lindahl, Myrdal, Ohlin, Lundberg)

Knut Wicksell llega a la economía desde las matemáticas y se forma leyendo a Böhm-Bawerk y a Walras, de modo que la escuela sueca nace ya como síntesis de ambas tradiciones. La contribución de Wicksell a la distribución se recoge en el apartado 2.5.1.

Interés y precios (1898) es el punto de partida de la principal tradición macroeconómica que se desarrolla dentro del periodo al margen de Cambridge y de Viena. Wicksell contrapone dos tipos de interés. El **tipo monetario** es el que cobra el sistema bancario por sus préstamos, mientras el **tipo de interés natural** refleja el rendimiento marginal del capital y es el que equilibra ahorro e inversión. Si los bancos prestan por debajo de ese rendimiento, la inversión desborda al ahorro y el exceso de demanda hace subir los precios en un proceso acumulativo, que no se detiene mientras se mantenga la diferencia. De aquí toma Hayek el armazón de la teoría austriaca del ciclo, y de aquí procede la idea de un tipo de interés de referencia no observable.

Sobre esa base, la escuela de Estocolmo (Lindahl, Myrdal, Ohlin y Lundberg) desarrolla en los años treinta dos instrumentos que anticipan el aparato keynesiano. La distinción entre magnitudes *ex ante* y *ex post* separa los planes de los agentes de los resultados realizados, y el análisis de secuencias sustituye la comparación de equilibrios por el seguimiento del ajuste en el tiempo. Ambos son anteriores e independientes respecto de la *Teoría general* (1936), lo que ha llevado a discutir la prioridad de Keynes en algunos de sus resultados.

La escuela quedó eclipsada por el éxito editorial de la *Teoría general* y por la barrera lingüística.

A4. Keynes: análisis comparado

El pensamiento económico de Keynes. Referencia a la economía postkeynesiana y el desequilibrio.

Introducción

Introducción común. Los temas A02 a A07 comparten una introducción, que se desarrolla al comienzo del tema **A02**. Recorre la evolución del pensamiento económico, del mercantilismo y la fisiocracia a la escuela clásica, la revolución marginalista, Keynes y la síntesis neoclásica, el monetarismo, la nueva macroeconomía clásica, la nueva economía keynesiana y la nueva síntesis neoclásica.

La Gran Depresión (1929-1933) mantuvo durante años un desempleo masivo en las economías industrializadas, algo que la teoría dominante no podía explicar. Para el **macromodelo neoclásico**, que reunía la ley de Say, la flexibilidad de precios y salarios y la teoría cuantitativa del dinero, el desempleo solo podía ser voluntario o friccional, y la economía volvía por sí sola al pleno empleo.

John Maynard Keynes respondió con la *Teoría general de la ocupación, el interés y el dinero* (1936), que rompe con ese marco. En ella, la **demanda efectiva** determina el nivel de producción y de empleo, y la economía puede quedar en un equilibrio con **desempleo involuntario**. El tipo de interés deja de ser el precio que iguala el ahorro y la inversión y pasa a depender de la **preferencia por la liquidez**, mientras que la incertidumbre que rodea a la inversión impide que la economía tienda por sí sola al pleno empleo. Un aumento del ahorro puede incluso reducir la renta (la **paradoja del ahorro**), y de ahí la defensa de una política pública que sostenga la demanda.

La recepción de la obra dio lugar a lecturas muy distintas. Hicks (1937) la tradujo al modelo IS-LM, que la síntesis neoclásica convirtió en el núcleo de la macroeconomía de posguerra, mientras que los poskeynesianos y los keynesianos del desequilibrio sostuvieron que esa traducción perdía lo esencial del mensaje, la incertidumbre y el ajuste por cantidades. El **efecto Pigou** (1943) planteó, por su parte, si la caída de los precios bastaba, a través de los saldos reales, para devolver la economía al pleno empleo. El debate sobre lo que Keynes dijo realmente, que reabrió Leijonhufvud (1968), sigue vivo, y la Gran Recesión y la pandemia devolvieron a primer plano las políticas de demanda de inspiración keynesiana.

Este tema estudia la aportación de John Maynard Keynes mediante la comparación de su macromodelo con el macromodelo neoclásico al que se opuso. Para ello, el tema se organiza en tres partes:

- *Macromodelo neoclásico*: el mercado de trabajo, el mercado de bienes con la ley de Say, el mercado de dinero con la teoría cuantitativa y el modelo de oferta y demanda agregadas.

- *Pensamiento de Keynes*, formalizado en el macromodelo keynesiano: la sustitución de paradigmas, el mercado de trabajo, el modelo IS-LM y el modelo de oferta y demanda agregadas.

- *Reacciones y desarrollos posteriores*: desde el debate sobre los saldos reales hasta la síntesis neoclásica, el poskeynesianismo y los keynesianos del desequilibrio.

1. Macromodelo neoclásico: una reconstrucción (Cassel, 1918b; Marshall, 1923; Fisher, 1930; Hicks, 1932; Pigou, 1933)

Los autores neoclásicos se centraron sobre todo en microfundamentar el pensamiento clásico, pero su teoría macroeconómica (teorías del ciclo, del crecimiento y monetarias) no estaba unificada antes de los años veinte. A partir de entonces, autores como John Hicks, Irving Fisher y Arthur Pigou se esforzaron por poner en contacto las teorías del ciclo con otras teorías económicas[1].

1.1. Principales rasgos

Metodología: la microfundamentación permitía aplicar modelos de equilibrio general al funcionamiento de la economía en su conjunto.

Supuestos:

- *Entorno*: los autores neoclásicos siguen la concepción clásica, "utópica", de las condiciones en que se desarrolla la actividad económica:
  - *Cumplimiento de la ley de Say*: toda oferta genera su propia demanda[2].
  - *Ausencia de fallos de mercado*: la información es perfecta y no existen poder de mercado, bienes públicos ni rigideces, de ahí que los precios y salarios sean perfectamente flexibles y todos los agentes, precio-aceptantes.
  - *Dicotomía clásica*: la economía real (oferta y demanda de bienes y servicios) puede modelizarse y tratarse por separado de la economía monetaria (oferta y demanda de dinero)[3].
- *Agentes*: son perfectamente racionales, en línea con la idea clásica del *homo economicus*.

Ajuste: de las características institucionales y la racionalidad de los agentes se deriva un mecanismo de ajuste walrasiano, vía precios plenamente flexibles (el salario real y el tipo de interés real).

Política económica: apoyan el no intervencionismo, porque las políticas de demanda no alteran la renta real ni el empleo (§1.5.2).

[1] El modelo, tal como se expone en esta sección, es una reconstrucción posterior que ninguno de los autores citados formuló completa. Keynes lo sistematizó en el capítulo 2 de la *Teoría general* (1936) como la teoría "clásica" que pretendía refutar, y Modigliani (1944) y Patinkin (1956) lo formalizaron después. Pigou (1933), por ejemplo, no negaba el paro, sino que lo atribuía a las resistencias que impiden ajustar los salarios al instante a los cambios de la demanda.

[2] La ley de Say no es la ley de Walras. Siguiendo a Lange (1942a), la ley de Say implica que la oferta es igual a la demanda en el conjunto de los mercados que no son monetarios (caben excesos parciales, no una sobreproducción general). La ley de Walras, por el contrario, es *a priori* aplicable a cualquier tipo de mercado. Por ello, la ley de Walras podría admitir un exceso de demanda u oferta en el sector monetario con un exceso compensador en el sector real. Sin embargo, esto no puede suceder bajo la ley de Say. Esa diferencia está en la base de la controversia Patinkin.

[3] De este modo, variables nominales (es decir, no solo el dinero, sino variables como los precios o los salarios nominales) se tratan por separado de las reales.

1.2. Mercado de trabajo

El modelo macroeconómico neoclásico es un **modelo de oferta**, porque, según la ley de Say, la causalidad de la actividad económica nace en la oferta de bienes y servicios. La oferta agregada se obtiene del mercado de trabajo, que determina el nivel de empleo, y de la función de producción agregada, que lo convierte en producto ofrecido (§1.5.1). Las dos curvas del mercado de trabajo, la oferta y la demanda, proceden de la extensión del análisis marginal a los salarios y el empleo, sobre todo en John Hicks (1932) y Arthur Pigou (1933).

1.2.1. Oferta de trabajo

El punto de partida es la decisión del consumidor-trabajador, que reparte su dotación de tiempo entre trabajo y ocio para maximizar su utilidad. En el óptimo iguala la tasa subjetiva a la que está dispuesto a intercambiar consumo y ocio (su relación marginal de sustitución) con el salario real, precio relativo de ambos bienes (el programa se recoge en el anexo, apartado A.1).

Si el efecto sustitución domina al efecto renta, la agregación, en sentido amplio, de las ofertas individuales de trabajadores homogéneos y con plena movilidad da una oferta de trabajo agregada creciente con el salario real:

$$L^s = L^s\left(\frac{w}{P}\right), \quad \frac{\partial L^s}{\partial (w/P)} > 0$$

1.2.2. Demanda de trabajo

Las empresas contratan trabajo solo como medio para producir bienes, de ahí que su demanda sea *derivada*. Con el *stock* de capital fijo del corto plazo, cada empresa competitiva maximiza su beneficio ampliando la plantilla mientras la productividad marginal del trabajo supere el salario real, hasta que ambos se igualan, $\partial f/\partial L = w/P$ (el programa se recoge en el anexo, apartado A.2).

De la agregación de las demandas individuales se obtiene una demanda de trabajo agregada decreciente con el salario real, de acuerdo con el principio de productividad marginal decreciente:

$$L^d = L^d\left(\frac{w}{P}\right), \quad \frac{\partial L^d}{\partial (w/P)} < 0$$

Como las empresas son salario-aceptantes, la causalidad va desde el salario hasta el empleo demandado. Esta dirección causal será uno de los puntos de ruptura de Keynes (§2.2.2).

> Los programas de elección consumo-ocio y de maximización del beneficio, con sus CPO y CSO, la estática comparativa y la evidencia empírica sobre las funciones de oferta y demanda de trabajo se desarrollan en el tema A25, apartados 1.2 y 1.3. La visión neoclásica del desempleo, en contraste con la keynesiana, se narra en el tema A28, apartado 1.1.

1.2.3. Equilibrio en el mercado de trabajo: el pleno empleo

Equilibrio: del cruce entre las curvas de oferta y demanda de trabajo agregadas (figura A4.1, panel del mercado de trabajo) se obtiene un equilibrio único, estable y eficiente (en el sentido paretiano), en el que la cantidad de trabajo de equilibrio equivale al nivel de pleno empleo de recursos de la economía. El salario real, que todos toman como dado, es un indicador perfecto de la escasez de trabajo.

Cuando se cumplen todos los supuestos neoclásicos, nadie está parado contra su voluntad. Solo quedan el paro friccional de búsqueda y los "no participantes", que dedican todo su tiempo al ocio. El *desempleo involuntario* (desempleo clásico), en cambio, requiere alguna rigidez que impida el ajuste salarial. Su causa es siempre un salario real demasiado alto, con el que la oferta de trabajo supera a la demanda.

Estática comparativa:

- *Políticas de demanda*: si la política fiscal o la monetaria mueven los precios, los salarios nominales varían en la misma proporción, así que las empresas mantienen sus planes de producción.

- *Perturbación de oferta*: un aumento de la productividad marginal del trabajo (progreso técnico o mayor *stock* de capital) desplaza la demanda de trabajo hacia la derecha y eleva el salario real y el empleo de equilibrio.

Dinámica: con salario nominal y precios perfectamente flexibles, cualquier perturbación se corrige automáticamente. Por ejemplo, un exceso de oferta de trabajo reduce el salario real hasta volver al equilibrio, y viceversa. La flexibilidad del salario real es así la primera línea de defensa de la no intervención (§1.5.2).

1.3. Mercado de bienes y servicios

La demanda real de bienes, o **gasto agregado**, está compuesta por el gasto privado en consumo, el gasto privado en inversión y el gasto público, todos ellos en términos reales y agregados:

$$Y^d = C(Y, r) + I(r) + G_0$$

donde $C(Y, r)$ denota la demanda de consumo, $I(r)$ la de inversión privada y G_0 el gasto público, exógeno como los ingresos públicos netos de transferencias, T_0.

Con la dicotomía clásica, este gasto real no fija ni la producción ni los precios. La producción la fijan el mercado de trabajo y la función de producción (§1.2 y §1.5.1), y el tipo de interés real se ajusta en el mercado de fondos prestables para que el gasto iguale esa producción a cualquier nivel de precios (§1.3.3 y §1.3.4). Por tanto, el gasto solo decide cómo se reparte la producción entre consumo, inversión y gasto público.

1.3.1. Demanda de consumo

Para los neoclásicos, el consumo no tiene más papel que cubrir las necesidades del momento y se considera poco más que un residuo del ahorro. Existen dos formas de construir la función de consumo agregada:

Fisher (1930): es la forma original, derivada del modelo de consumo intertemporal en dos periodos. El consumo real agregado es una función creciente de la renta y, si el efecto sustitución

domina al efecto renta, decreciente del tipo de interés real, $C = C(Y^+, r^-)$. Correlativamente, el ahorro es creciente en el tipo de interés real, $S_F = S_F(r^+)$. Con esta formulación se cumplen la dicotomía clásica y la neutralidad del dinero (el agregado monetario no forma parte de la riqueza).

Pigou (1943): si se considera, además, el efecto saldos reales o *efecto Pigou* (§3.1.2), el consumo depende también de la riqueza real neta, $C = C(Y^+, r^-, W^+)$.

1.3.2. Demanda de inversión

Nadie atesora dinero, así que todo el ahorro acaba convertido en inversión. En la "segunda aproximación" de Irving Fisher, la tasa de retorno marginal sobre el coste se iguala al tipo de interés real, de modo que, con productividad marginal del capital decreciente, cuanto mayor es el tipo de interés, menor es la inversión que resulta rentable. De aquí se obtiene una función de demanda de inversión agregada decreciente en el tipo de interés real:

$$I = I(r), \quad \frac{\partial I}{\partial r} < 0$$

> La derivación de la función de consumo neoclásica a partir del modelo intertemporal de Fisher se desarrolla en el tema A33, apartados 1.1 y 2.1. La de la demanda de inversión (tasa de retorno sobre el coste) y el mercado de fondos prestables, con los efectos de las políticas fiscal y monetaria, se encuentra en el tema A34, apartado 1.1.

1.3.3. Equilibrio: la ley de Say

Con base en estas consideraciones se puede obtener el equilibrio en el mercado de bienes y servicios finales[4]:

$$Y^d = C(Y, r) + I(r) + G_0 \tag{1}$$

$$Y = C(Y, r) + S_F(r) + T_0 \tag{2}$$

$$Y^d = Y \;\rightarrow\; I = S_F + (T_0 - G_0) = S_F + S_G \tag{3}$$

$$S = I \tag{4}$$

La ecuación (1) recoge el comportamiento del gasto y (2) es la identidad que reparte la renta entre consumo, ahorro e impuestos. La condición de equilibrio es (3). Si la demanda de bienes coincide con la oferta de pleno empleo, el mercado se vacía y la inversión se financia con el ahorro privado y el público, $S_G = T_0 - G_0$. Por último, (4) reescribe (3) como igualdad entre ahorro nacional e inversión nacional.

[4] La igualdad entre ahorro e inversión de las ecuaciones (3) y (4) se cumple siempre *ex post*, como identidad contable, pero aquí se lee como condición de equilibrio entre magnitudes planeadas (*ex ante*), de ahí el signo $=$ y no $\equiv$.

Las características institucionales del mercado de trabajo garantizan el pleno empleo, y la ley de Say, el vaciado del mercado de bienes. Por ello, clásicos y neoclásicos sostenían que la renta de equilibrio, a corto y a largo plazo, era la de pleno empleo.

1.3.4. Ajuste en el mercado de bienes (teoría no monetaria del interés): la flexibilidad del tipo de interés natural

Si la flexibilidad del salario real era la primera línea de defensa de la tesis del pleno empleo, la segunda es la **flexibilidad del tipo de interés real**, que garantiza que cambios exógenos en el gasto agregado no perturban el mercado de bienes y servicios.

Origen: para explicar cómo se determina el tipo de interés real, varios autores, de Mill (1848) a Fisher (1907), plantean una teoría no monetaria del interés basada en el mercado de fondos prestables[5].

Idea: en ese mercado concurren el ahorro de los consumidores, la inversión de las empresas y el déficit del sector público, y el tipo de interés real[6] equilibra ese intercambio, que se canaliza mediante bonos o promesas de pago futuras.

Planteamiento: el modelo consta de dos funciones:

- *Oferta de fondos prestables (demanda de bonos)*: es el ahorro privado, $S_F(r^+)$. El tipo de interés real remunera, por tanto, la espera, es decir, el aplazamiento del consumo.

- *Demanda de fondos prestables (oferta de bonos)*: es la inversión privada, $I(r^-)$, más el déficit público exógeno, $(G_0 - T_0)$, que el sector público financia mediante la emisión de deuda.

Equilibrio: de la interacción de ambas se obtiene el tipo de interés real de equilibrio, que reescribe la condición (3) en términos del déficit público:

$$S_F(\bar{r}) = I(\bar{r}) + (G_0 - T_0)$$

El equilibrio es único y estable. El tipo de interés de equilibrio, $\bar{r}$, es el *tipo de interés natural*, término acuñado por Knut Wicksell (1898). Se determina por fuerzas reales (el ahorro y la productividad marginal del capital), es compatible con la estabilidad de precios y es el valor hacia el que gravita a largo plazo el tipo de interés monetario, i, que se determina en los mercados monetarios.

Estática comparativa: el tipo de interés natural solo varía con las perturbaciones reales que desplazan la oferta o la demanda de fondos prestables (véase la figura del tema A34, apartado 1.1.2):

- *Política fiscal*: un aumento del déficit público eleva la demanda de fondos prestables y el tipo de interés natural, y expulsa inversión privada y consumo (§1.5.2).

[5] A menudo, la expresión "teoría de los fondos prestables" designa indistintamente dos teorías. Una es la teoría no monetaria sobre el tipo de interés debida, principalmente, a las aportaciones de John Stuart Mill (1848) e Irving Fisher (1907), en la que $S = I$. La otra es la teoría de los fondos prestables en sí misma, formulada por Dennis Robertson (1926) y Bertil Ohlin (1937), con antecedentes en Knut Wicksell (1898), en la que entra en juego el crédito y no necesariamente $S = I$.

[6] El tipo de interés real que reciben los ahorradores es $r = i - \pi^e$, donde i es el tipo de interés nominal de los bonos y π^e es la tasa de inflación esperada.

▪ *Perturbaciones reales*: un aumento de la propensión al ahorro reduce el tipo de interés natural y eleva la inversión, mientras que una mejora de la productividad marginal del capital lo eleva, junto con el ahorro y la inversión.

Conclusiones:

▪ *Aspectos positivos*: ofrece una explicación real del tipo de interés, ligada al ahorro y a la productividad del capital, y su concepto de tipo de interés natural pervive como referencia de la política monetaria en la nueva economía keynesiana.

▪ *Limitaciones*: Keynes le objetó que el ahorro depende sobre todo de la renta, de modo que su igualdad con la inversión se alcanza mediante variaciones de la renta y no del tipo de interés. Le objetó también que el tipo de interés es un fenómeno monetario, la recompensa por renunciar a la liquidez (§2.3.3).

1.4. Mercado de dinero: la teoría cuantitativa (Fisher y Cambridge)

Por la dicotomía clásica, el análisis anterior se refiere solo a la economía real, mientras que la economía monetaria se explica a través de la teoría cuantitativa del dinero.

Origen: la relación proporcional entre dinero y precios había sido ya anunciada por M. de Azpilcueta y la escuela de Salamanca (1556)[7], J. Bodin (1568) y, con mayor difusión, D. Hume (1752). Su análisis formal corresponde a dos corrientes independientes, el enfoque de Fisher (1911), macroeconómico y centrado en los factores institucionales, y el de Cambridge (1917), microeconómico y centrado en las decisiones individuales.

Idea: la teoría cuantitativa es una teoría sobre las causas de los cambios del valor del dinero, cuya conclusión es que, a largo plazo, el nivel general de precios guarda una relación proporcional con la cantidad de dinero. En ambos enfoques la oferta de dinero es exógena y determinada por las autoridades monetarias.

Razonamiento: Fisher reinterpreta causalmente, de izquierda a derecha, la identidad contable del gasto (la "ecuación de cambio"), y deja el nivel de precios como única incógnita del modelo. La escuela de Cambridge, por su parte, llega a la "ecuación de Cambridge" a partir de la idea de que los agentes desean mantener una fracción k de su renta en forma de saldos de caja:

$$MV = PY \qquad \Longleftrightarrow \qquad M^d = kPY, \quad k = \frac{1}{V}$$

La demanda de dinero es "mecanicista" y muy estable, ya que el dinero solo se usa como medio de cambio. Así, V (o la k de Cambridge) queda determinada por factores institucionales y no por el tipo de interés, si bien el enfoque de saldos de caja admite que este pueda afectar a la demanda de dinero.

Se cumple, por tanto, la *neutralidad del dinero*. Como la renta se determina en el mercado de trabajo y V es estable, el dinero es un velo, y sus variaciones solo mueven en la misma proporción los precios y los salarios nominales.

> La derivación de los enfoques de las transacciones (Newcomb-Fisher) y de los saldos de caja (Cambridge) se desarrolla en el tema A35, apartado 1.1, y su lectura como teoría neoclásica de la inflación, en el tema A41, apartado 2.1.1.

[7] M. de Azpilcueta desarrolló una teoría del valor-escasez precursora de la teoría cuantitativa del dinero en su *Comentario resolutorio de cambios* (1556).

Conclusiones:

- *Aspectos positivos*: explica la relación de largo plazo entre la cantidad de dinero y el nivel de precios, que la evidencia confirma sobre todo en los países con inflación elevada (De Grauwe y Polan, 2005), y es la base de la reformulación monetarista de Friedman.

- *Limitaciones*: trata la velocidad de circulación como un dato institucional, cuando la demanda de dinero depende también del tipo de interés, como mostrará la preferencia por la liquidez de Keynes (§2.3.3), y en la práctica la velocidad ha resultado inestable.

1.5. Macromodelo neoclásico OA-DA

La aportación neoclásica se resume en un modelo de oferta y demanda agregada con precios flexibles. La oferta agregada, derivada del mercado de trabajo, fija la producción, y la demanda agregada, derivada de la teoría cuantitativa, solo el nivel de precios. Esta demanda agregada no es el gasto real del §1.3. Como, con la dicotomía clásica, el tipo de interés iguala ese gasto con $\bar{Y}$ a cualquier nivel de precios, lo único que liga el gasto con los precios es el dinero. La DA recoge así las combinaciones de renta y precios compatibles con la ecuación de cambio, $PY = MV$, para M y V dados.

1.5.1. Oferta y demanda agregada

Oferta agregada: al sustituir el nivel de empleo de equilibrio, L^* (§1.2.3), en la función de producción agregada, con el *stock* de capital $\bar{K}$ dado en el corto plazo, se obtiene la oferta agregada de bienes[8]:

$$Y = f(\bar{K}, L) \;\rightarrow\; Y^s = f(\bar{K}, L^*)$$

La oferta agregada neoclásica es así totalmente rígida al tipo de interés real y a los precios, y vertical en los planos renta-tipo de interés y renta-precios.

Desarrollo gráfico: el macromodelo neoclásico se representa en el plano renta real-precios, con los paneles auxiliares del mercado de trabajo, la función de producción y la recta de 45° (figura A4.1):

1. *Mercado de trabajo*: en el plano empleo-salario real, el cruce de la demanda de trabajo, L^d, y la oferta, L^s, fija el pleno empleo, L^*, y el salario real $(w/P)^*$.

2. *Función de producción*, $f(L)$: traslada L^* a la producción de pleno empleo, $\bar{Y}$.

3. *Recta de 45°*: lleva $\bar{Y}$ al eje horizontal del plano renta real-precios.

4. *Oferta agregada*, OA: vertical en $\bar{Y}$, porque la producción no depende de los precios.

5. *Demanda agregada*, DA: procede de la ecuación de cambio (§1.4), $Y = MV/P$, y es decreciente porque, con MV dado, un nivel de precios más alto compra menos producción.

6. *Equilibrio*: en el corte de la OA y la DA, la renta es $\bar{Y}$ y la demanda agregada solo fija el nivel de precios.

[8] A diferencia de la oferta de mercado microeconómica, que suma las ofertas individuales de empresas que maximizan el beneficio a precios dados, la oferta agregada neoclásica se obtiene del equilibrio del mercado de trabajo.

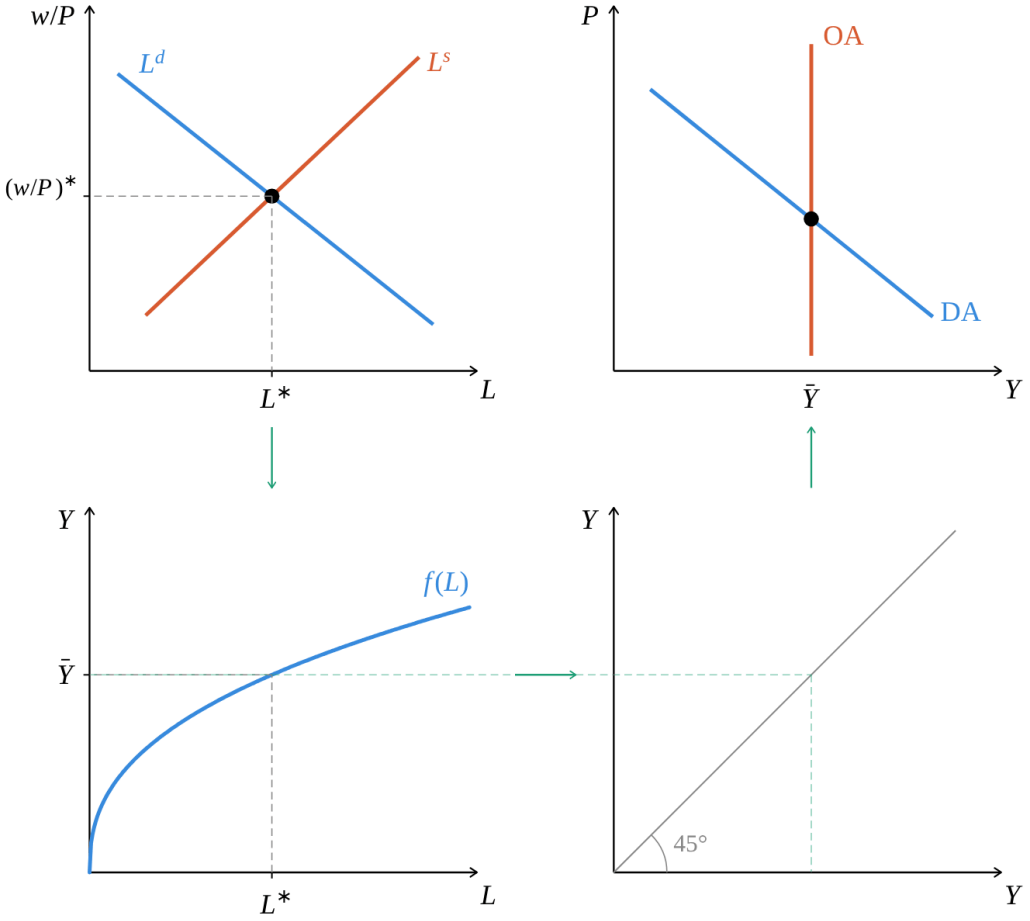

Figura A4.1: Macromodelo neoclásico. El mercado de trabajo determina el empleo y la producción de pleno empleo. La oferta agregada es vertical y la demanda agregada (teoría cuantitativa) solo fija el nivel de precios.

1.5.2. Implicaciones de política económica

En el paradigma neoclásico, las políticas de demanda (fiscales o monetarias) no influyen en la renta real ni en el empleo, y las de oferta sí pueden hacerlo:

- *Política monetaria*: una expansión monetaria puede reducir transitoriamente el tipo de interés de mercado por debajo del natural (el proceso acumulativo de Wicksell). Sin embargo, el exceso de demanda de bienes resultante se corrige con una subida proporcional de precios y salarios nominales, sin efectos reales (neutralidad del dinero) y sin alterar el tipo de interés natural (§1.3.4). Su papel debe ceñirse, por tanto, a dar liquidez a la economía y contener la inflación.

- *Política fiscal*: una expansión fiscal obliga a captar más recursos en el mercado de fondos prestables, eleva así el tipo de interés real y expulsa consumo e inversión privados por un importe idéntico al aumento del gasto público, con un efecto expulsión (*crowding out*) total[9]. No afecta

[9]Esta conclusión tuvo su versión práctica en la *Treasury view* británica, que el ministro de Hacienda, Winston Churchill, expuso en su discurso presupuestario de abril de 1929. Según ella, el gasto público financiado con deuda compite con la industria por los mismos recursos y apenas crea empleo adicional. Keynes y Hubert

a la renta real ni al empleo, solo al tipo de interés real y a la composición del gasto, y, con la velocidad de circulación fijada por factores institucionales (§1.4), tampoco al nivel de precios. Por ello debe cumplirse la regla del presupuesto equilibrado en todo momento[10].

▪ *Políticas de oferta*: la intervención solo se justifica para eliminar los obstáculos institucionales que impidan cumplir los supuestos del modelo (p. ej. rigideces salariales). En cambio, los tributos sobre el trabajo, las subvenciones o los programas de formación sí afectan a la renta real y al empleo.

1.6. Conclusiones

Aspectos positivos: los esfuerzos de autores neoclásicos como Cassel, Marshall, Pigou o Fisher permitieron integrar muchas de las relaciones en un primer intento de modelar el ámbito agregado de la economía.

Limitaciones:

▪ *Ley de Say*: es un eslabón débil del modelo. Solo resulta plenamente válida en una economía de trueque, porque en una economía monetaria exige suponer que el dinero sirve únicamente como medio de pago y que nadie lo atesora.

▪ *Dicotomía*: es poco plausible considerar una independencia estricta (es decir, a corto y largo plazo) entre el sector real y el monetario.

▪ *Política económica*: la defensa del *laissez faire* (p. ej. la regla del presupuesto equilibrado) puede no ser aconsejable.

Henderson la rebatieron ese mismo año en *Can Lloyd George Do It?*, y el multiplicador de Kahn (1931) dio después base teórica a su réplica.

[10] A lo sumo, se admitiría el endeudamiento para financiar inversiones autoliquidables, que es lo que hoy suele denominarse "regla de oro" de las finanzas públicas.

2. Pensamiento de Keynes: macromodelo keynesiano (Keynes, 1936; Hicks, 1937; Modigliani, 1944; Hansen, 1953)

Las predicciones neoclásicas no se ajustaron a la evidencia de la Gran Depresión (1929-1933) y su elevado desempleo involuntario, que está en el origen de la obra de Keynes. Su *Teoría general de la ocupación, el interés y el dinero* (1936) supone una sustitución de paradigmas, en términos de Kuhn, y sienta las bases de la escuela keynesiana. En su *Tratado sobre el dinero* (1930), los desequilibrios entre ahorro e inversión, originados por la divergencia entre el tipo de interés de mercado y el natural, se ajustaban vía nivel de precios y beneficios extraordinarios. En la *Teoría general*, en cambio, la principal variable de ajuste es la producción y, con ella, la renta.

Keynes no presentó su teoría como un modelo formal. Por eso se expone con el modelo renta-gasto con mercado de trabajo (IS-LM con mercado de trabajo), en su versión con microfundamentación puramente keynesiana, que resume con sencillez la interacción entre la economía real y la monetaria.

El modelo de la síntesis neoclásica, cuya formación se narra en el tema A05, apartado 1.1, reúne el diagrama IS-LM de Hicks (*Mr. Keynes and the Classics*, 1937)[11], que Alvin Hansen (1953) difundió en Estados Unidos, el mercado de trabajo con salario nominal rígido de Modigliani (1944) y el aspa keynesiana de Samuelson (1948a). El modelo de la síntesis perseguía formalizar las ideas de Keynes y aunarlas con las neoclásicas (§2.4.3).

Los supuestos iniciales son corto plazo, economía cerrada y cuatro mercados (trabajo, bienes, dinero y bonos), de los que el de bonos queda implícito por la ley de Walras (§2.3.3). Como primera aproximación se suponen precios fijos, un supuesto que se levanta en el §2.4. Con ellos, la oferta agregada se adapta pasivamente a la demanda y el tipo de interés real coincide con el nominal, $r = i$.

2.1. Principales rasgos: la sustitución de paradigmas

Metodología: Keynes combina el razonamiento inductivo con el deductivo en una teoría "realítica" (realista y analítica), en la expresión de Landreth y Colander.

Supuestos:

- *Entorno*: presenta importantes diferencias con la concepción clásica:

  - *Ley de Say*: Keynes invierte la relación causal clásica entre oferta y demanda, y la *demanda efectiva*, es decir, los ingresos que los empresarios esperan obtener de sus ventas, pasa a determinar la producción y el empleo. Una insuficiencia de demanda efectiva restringe así la producción a corto plazo y puede dar lugar a desempleo involuntario.

  - *Incertidumbre radical*: la imposibilidad de reducir el futuro a un riesgo cuantificable probabilísticamente hace que las expectativas desempeñen un papel central (si bien Keynes no ofrece un modelo de formación de estas).

[11] Hicks llamó LL a la curva de equilibrio monetario, y el nombre LM se generalizó después. Su artículo procede de la ponencia que presentó en la reunión de la Econometric Society celebrada en Oxford en septiembre de 1936, en la que Roy Harrod y James Meade expusieron sistemas de ecuaciones parecidos para interpretar la *Teoría general*.

- *Rigidez salarial nominal a la baja*: convive con precios flexibles (competencia perfecta en el mercado de bienes), y Keynes la atribuye a la preocupación de los trabajadores por su salario relativo (§2.2.1). Además, sostiene que, aunque los salarios fueran flexibles, su bajada no garantizaría el pleno empleo (§3.1.1).

- *No dicotomía*: Keynes no distingue de forma explícita entre economía monetaria y economía real.

▪ *Agentes*: como los neoclásicos, Keynes cree que los individuos se mueven por el interés personal, pero deciden bajo incertidumbre radical y se apoyan en convenciones y en los *animal spirits*. La lectura habitual de su mercado de trabajo añade la *ilusión monetaria* de los trabajadores (§2.2.1)[12].

Ajuste: a diferencia del modelo neoclásico, Keynes se inspira en Malthus y Marshall y considera un *mecanismo marshalliano*, es decir, de ajuste vía cantidades. Las variables de ajuste son las existencias empresariales (la inversión no planeada) y, a través de ellas, la producción y la renta.

Política económica: aboga por un mayor intervencionismo, sobre todo por el lado de la demanda (§2.4.2).

2.2. Mercado de alquiler de trabajo

Keynes se opone al modelo neoclásico del mercado de trabajo, siempre en pleno empleo, y concibe la posibilidad de desempleo involuntario[13] incluso cuando existe competencia perfecta en el mercado de bienes y servicios[14]. Para Keynes, los salarios no solo son un coste para las empresas, sino renta de los trabajadores, de modo que la demanda agregada desempeña un papel esencial en el mercado de trabajo.

2.2.1. Oferta de trabajo

Idea: la *Teoría general* prescinde de una curva de oferta de trabajo. Su argumento suele leerse en clave de ilusión monetaria, según la cual los trabajadores fijan sus exigencias en salario nominal y no en salario real, es decir, maximizan su utilidad sin atender al nivel de precios[15].

Razonamiento: la ilusión monetaria tiene dos implicaciones en función de si suben o bajan los precios:

▪ *Bajada de precios*: los trabajadores ven en cualquier rebaja del salario nominal una merma de su poder adquisitivo y de su posición relativa, y se oponen a ella. Por ello, el salario nominal presenta rigidez a la baja y en el tiempo, con un mínimo, w_m, fijado por factores institucionales.

[12] Fisher dedicó al concepto de ilusión monetaria su libro *The Money Illusion* (1928). Su relación con la visión keynesiana del mercado de trabajo se estudia en el tema A28, apartado 1.1.2.

[13] Keynes define el desempleo involuntario en el capítulo 2 de la *Teoría general* con un criterio operativo. Hay paro involuntario si, ante una pequeña subida del precio de los bienes que consumen los asalariados en relación con el salario nominal, tanto el número de trabajadores dispuestos a trabajar por el salario nominal vigente como la demanda de trabajo a ese salario superan el empleo existente. Los trabajadores aceptarían, por tanto, una bajada del salario real que llegara por la vía de los precios, aunque se resistan a la de su salario nominal (§2.2.1).

[14] Robinson ya había criticado la teoría neoclásica al mostrar que, en monopsonio, el empleo y el salario quedan por debajo de los competitivos.

[15] Keynes rechaza así el segundo postulado clásico: el salario real no tiene por qué igualar la desutilidad marginal del trabajo.

- *Subida de precios*: los trabajadores negocian sin reparar en el alza de precios, así que aceptan un salario real menor y se avienen a trabajar más horas[16]. Esta conducta asimétrica solo es irracional en apariencia, pues se explica por el salario relativo. Cada trabajador vigila su lugar en la escala salarial, y una subida generalizada de los precios, al afectar a todos por igual, deja intacto su salario real relativo.

Por ello, aunque precios y tipo de interés sean flexibles y haya competencia perfecta, el pleno empleo no está garantizado, y puede alcanzarse sin esas condiciones[17]. Con todo, tampoco la rigidez salarial nominal es condición necesaria ni suficiente para la existencia de desempleo involuntario:

1. *No necesaria*: también puede existir desempleo involuntario por causa de los dos eslabones débiles de Keynes (apartados 2.3.2 y 2.3.3) u otras rigideces.

2. *No suficiente*: con demanda agregada suficiente, el desempleo es voluntario (§2.2.3).

Formalmente, la oferta de trabajo agregada se puede definir como una función a trozos del salario ofrecido:

$$L^s = L^s(w), \qquad \begin{cases} L^s = 0 & \forall w < w_m \\ L^s > 0 & \forall w \geq w_m \end{cases}$$

donde w es el salario de mercado y w_m el suelo del salario nominal por debajo del cual los agentes no ofrecen trabajo. No debe confundirse con un salario de reserva, porque no refleja la valoración individual del ocio, sino una norma institucional sobre el salario nominal.

Desarrollo gráfico: en el plano empleo-salario nominal[18] (figura A4.2), la oferta de trabajo agregada presenta tres tramos:

1. *Tramo horizontal*: es perfectamente elástico en el nivel de salario mínimo, porque el desempleo involuntario es tan alto que los trabajadores están dispuestos a trabajar a ese salario (aunque no a menos).

2. *Tramo creciente*: aparece conforme la economía se reactiva y los sindicatos ganan fuerza.

3. *Tramo rígido*: corresponde al momento en que se alcanza el pleno empleo de recursos.

Si el equilibrio en el mercado de trabajo se da en el tramo horizontal (punto 0), el desempleo involuntario, U_I, es la diferencia entre el empleo de equilibrio al salario mínimo, L^*, y el nivel de empleo a partir del cual la oferta de trabajo es creciente, L_1. Entre L_1 y $\bar{L}$, en cambio, el paro es voluntario, U_V, porque esos trabajadores solo ofrecerían su trabajo a un salario mayor.

[16] Esto se debe a que, en el plano empleo-salario nominal, el aumento de precios desplaza hacia la derecha la demanda de trabajo y provoca un movimiento a lo largo de la oferta de trabajo.

[17] Keynes va más allá y sostiene que bajar precios y salarios podría incluso agravar la situación. Con la deflación crece el peso real de las deudas, la situación financiera de las empresas se deteriora y la inversión cae (efecto distribución). Además, cada caída de precios y salarios alimenta la creencia de que vendrán otras, lo que lleva a hogares y empresas a aplazar su gasto (efecto expectativas).

[18] Existen representaciones alternativas con el salario real, y la elección depende de la especificación de la oferta agregada. Se sigue aquí la exposición de Bajo y Monés (2000).

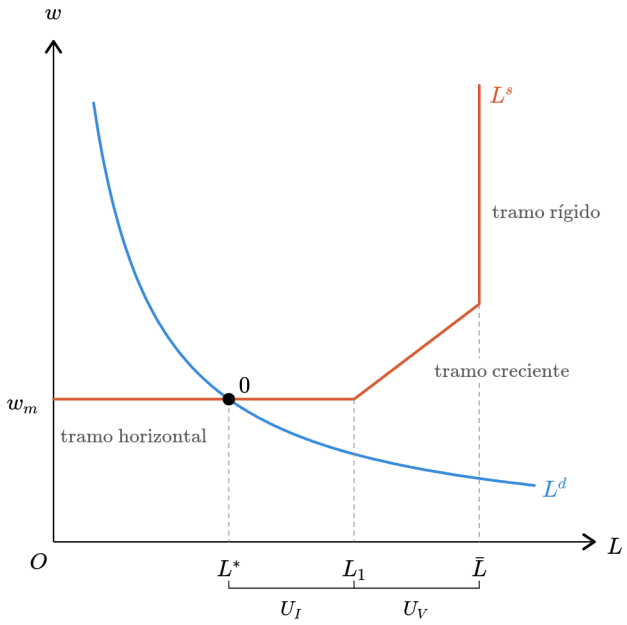

Figura A4.2: Oferta de trabajo keynesiana. Con el salario nominal rígido a la baja en w_m, la oferta es horizontal hasta L_1, creciente hasta el pleno empleo y vertical en $\bar{L}$. En el equilibrio 0 hay desempleo involuntario, U_I, y, entre L_1 y $\bar{L}$, voluntario, U_V.

Conclusiones:

▪ *Aspectos positivos*: al impedir que el salario nominal baje, hace depender el empleo de la demanda agregada, y fundamenta esa rigidez en la preocupación por el salario relativo, sin atribuir a los trabajadores una conducta irracional.

▪ *Limitaciones*: justifica la rigidez del salario nominal de forma informal, sin un modelo de la conducta de trabajadores y empresas que la explique. Además, la lectura habitual de esta oferta de trabajo en términos de ilusión monetaria no resiste la incorporación de las expectativas de inflación, como mostrará la crítica de Friedman y Phelps (tema A05, apartado 2.3.2).

2.2.2. Demanda de trabajo

En la demanda de trabajo Keynes sigue a los neoclásicos, con la condición de contratar hasta que $\partial f/\partial L = w/P$, que en el plano (L, w) da una curva decreciente. La única diferencia es que la causalidad va ahora desde el empleo hasta el salario, de modo que una demanda insuficiente de bienes y servicios puede originar desempleo involuntario.

2.2.3. Equilibrio

El equilibrio se obtiene de la intersección de ambas curvas en el plano empleo-salario nominal (figura A4.2). A diferencia del mercado de trabajo neoclásico, caben dos situaciones:

1. *Insuficiencia de demanda agregada*: el equilibrio en el mercado de trabajo se da en el tramo horizontal y hay desempleo involuntario.

2. *Suficiencia de la demanda agregada*: el equilibrio se da en el tramo creciente y el desempleo es voluntario, aunque todavía existen recursos ociosos (§2.4.1).

2.2.4. Implicaciones de política económica

1. *Políticas de demanda*: las expansivas son fundamentales frente a la insuficiencia de demanda agregada y el desempleo involuntario, siempre que existan recursos ociosos.

2. *Políticas de oferta*: Keynes no apoya las reducciones del salario nominal, pues a su juicio solo elevarían el empleo de forma indirecta, a través del efecto Keynes (§3.1.1), al igual que una expansión monetaria.

2.3. Mercado de bienes, servicios y dinero: modelo IS-LM

Como la existencia de desempleo involuntario depende de la demanda agregada, el modelo keynesiano está dominado por el lado de la demanda. Además, como Keynes rompe la dicotomía clásica, el mercado de bienes y servicios se trata conjuntamente con el de dinero.

2.3.1. Oferta agregada de bienes y servicios. Oferta agregada de dinero

1. *Oferta de bienes*: como inicialmente se suponen precios fijos, el análisis se centra por ahora en el lado de la demanda.

2. *Oferta de dinero*: Keynes la considera exógena, al igual que los neoclásicos (excepto Wicksell).

2.3.2. Demanda agregada: equilibrio en el mercado de bienes y servicios. Curva IS

La demanda agregada se obtiene de dos relaciones, la IS, con tres componentes (consumo, inversión y gasto público reales), y la LM, que recoge el mercado de dinero.

Consumo: frente a su tratamiento residual en el paradigma clásico, el consumo es para Keynes el fin último de la actividad económica y la base del proceso multiplicador. Depende de la renta corriente, y no del tipo de interés, y su relación con ella obedece a una "ley psicológica fundamental", según la cual, ante una unidad adicional de renta, los consumidores aumentan su consumo, pero en una cantidad menor.

La demanda agregada de consumo es así una función lineal y estable de la renta real absoluta corriente, con un componente autónomo (consumo de subsistencia) y un componente endógeno gobernado por la *propensión marginal a consumir*[19]:

$$C = C_0 + cY, \quad C_0 > 0,\ c \in (0,1)$$

El ahorro es entonces un excedente de ingreso, muy poco influido por el tipo de interés. De aquí se deducen dos implicaciones:

[19] Keynes (1936), en traducción propia: "la propensión a consumir es una función razonablemente estable, de modo que, por regla general, la cuantía del consumo agregado depende sobre todo de la cuantía de la renta agregada".

▪ *Multiplicador*: un aumento del gasto autónomo eleva la renta en un múltiplo, $1/(1-c) > 1$, de dicho aumento. El multiplicador keynesiano[20] muestra que la causalidad va desde la demanda hacia la oferta. Por ello, la política fiscal será muy eficaz sobre la renta, tanto más cuanto mayor sea c, y las políticas redistributivas estimularán el consumo, ya que los hogares de menor renta tienen mayor propensión marginal a consumir.

▪ *Paradoja del ahorro*: un aumento del ahorro autónomo, *ceteris paribus*, reduce la demanda agregada y, vía multiplicador, la renta, de modo que no eleva necesariamente el ahorro agregado y puede obstaculizar el pleno empleo. La paradoja ataca así la defensa clásica del ahorro como fuente del crecimiento.

Las primeras contrastaciones empíricas parecían corroborar la función de consumo keynesiana, pero la evidencia de Simon Kuznets, según la cual la propensión media al consumo es constante a largo plazo[21], obligó a considerar la renta a largo plazo y la riqueza. De ello se ocuparon la renta relativa y el ciclo vital de la síntesis neoclásica (§3.2) y la renta permanente de Friedman (tema A05, apartado 2.2.1).

Inversión: Keynes distingue la inversión *planeada*, componente de la demanda agregada, de la *efectiva*, que se integra en la renta nacional y difiere de la primera en la variación no planeada de existencias (§2.3.5). Para invertir la ley de Say, Keynes invierte también la relación entre ahorro e inversión, de modo que la inversión, movida por las expectativas de los empresarios, determina el ahorro.

Quien invierte compra un derecho sobre un flujo de retornos esperados, de modo que, a diferencia del enfoque fisheriano, las expectativas pesan en la decisión. Se invierte mientras la *eficiencia marginal del capital* (EMK), tasa de descuento que iguala el valor actual de los rendimientos esperados con el coste de reemplazo del bien, supere al tipo de interés, coste de oportunidad de la inversión. Como la EMK decrece con el volumen de inversión (caen los rendimientos esperados y sube el coste de reemplazo de los bienes de capital), el óptimo se alcanza cuando EMK $= r$. De aquí se obtiene una función de demanda de inversión agregada[22]:

$$I = I(r) = I_0 - vr, \quad v > 0$$

donde I_0 es el componente exógeno, que refleja el peso de las expectativas o *animal spirits* y toma un valor elevado, y v mide la sensibilidad de la inversión al tipo de interés, que se considera baja. Esa baja sensibilidad, primer eslabón débil de Keynes, es, junto con la trampa de liquidez (segundo eslabón débil) y la ilusión monetaria, una de las causas del desempleo involuntario.

Expectativas a largo plazo: Keynes (1936) distingue las expectativas a corto plazo, sobre los ingresos de la producción en curso, que guían las decisiones de producción y empleo, de las expectativas a largo plazo, sobre los rendimientos futuros de los bienes de capital, que guían la inversión. Como apenas se sabe nada de los rendimientos lejanos, estas últimas dependen del *estado de la confianza* y descansan en una *convención*, la de suponer que la situación presente continuará salvo que haya razones concretas para esperar un cambio. La bolsa, que revalúa a diario las inversiones, hace frágil esa convención, porque el especulador no estima el rendimiento

[20] Se le llama keynesiano, pero, en realidad, la idea se debe al colega de Keynes en Cambridge, Richard Kahn (1931).

[21] Kuznets (1946), con series de renta y consumo de EE. UU. de 1869 a 1938. La función keynesiana implica una propensión media decreciente, $C_t/Y_t = C_0/Y_t + c$, con $\frac{d}{dY_t}(C_t/Y_t) = -C_0/Y_t^2 < 0$, mientras que a largo plazo se observa una propensión media constante, $C_t/Y_t = \bar{c}$, mayor que la propensión marginal de corto plazo, c.

[22] La inversión es un gasto independiente del nivel de ingreso, ya que las decisiones de inversión, aunque generan una serie de compromisos a corto plazo, se toman a largo plazo. Por ello, dependen de las expectativas psicológicas a largo plazo y quedan al margen de las condiciones de ingreso.

de los activos, sino lo que la opinión media espera que espere la opinión media, como en un *concurso de belleza*. De ahí que la EMK sea inestable y que I_0 se desplace con el ánimo de los inversores. La misma raíz convencional tiene el tipo de interés "normal" de la demanda especulativa de dinero (§2.3.3), de modo que la inversión y la preferencia por la liquidez comparten el origen de la inestabilidad keynesiana.

Gasto público: Keynes lo considera una variable exógena, G_0, al igual que los clásicos, si bien con un papel muy diferente (§2.4.2).

Curva IS: agrupando todos los componentes del gasto, la relación IS muestra las combinaciones renta-tipo de interés que mantienen el equilibrio en el mercado de bienes, $Y = C + I + G_0$:

$$Y = \frac{1}{1-c}\left[C_0 + I_0 + G_0 - vr\right] = \frac{1}{1-c}\left[A_0 - vr\right]$$

donde el primer término es el multiplicador y A_0 agrupa los componentes autónomos del gasto. Gráficamente, en el plano renta-tipo de interés, la IS es una recta decreciente (véase la figura del tema A05, apartado 1.2.1), ya que al disminuir el tipo de interés aumenta la inversión y, por el multiplicador, la renta de equilibrio. Cuanto menor sea v, más inclinada será la IS, y lo es más para los keynesianos que para los monetaristas. Por su parte, los componentes autónomos desplazan la curva.

> La función de consumo keynesiana con impuestos, los multiplicadores fiscales, la paradoja del ahorro y la evidencia de Kuznets se desarrollan en el tema A33, apartado 1.2.1, y la eficiencia marginal del capital y su desarrollo gráfico, en el tema A34, apartado 1.2. La derivación de la relación IS, su estática comparativa y su análisis gráfico se encuentran en el tema A05, apartado 1.2.1, y los efectos de la política fiscal en el IS-LM, en el tema A38, apartado 1.1.

2.3.3. Demanda agregada: teoría de la preferencia por la liquidez y el equilibrio monetario. Curva LM

Teoría de la preferencia por la liquidez: en su teoría monetaria, Keynes solo acepta la teoría cuantitativa como caso límite, pues, bajo supuestos simplificadores, mientras haya desempleo una variación de la cantidad de dinero altera sobre todo el empleo. Solo en pleno empleo los precios varían en la misma proporción que la cantidad de dinero. Keynes enfatiza, además, el papel de la incertidumbre y las expectativas en el corto plazo, y define tres motivos por los que se demanda dinero:

1. *Transacción*: para gastos planeados, por la función del dinero como medio de cambio. Keynes acepta aquí el enfoque de Cambridge, con una demanda real proporcional a la renta real, $M_T^d = M_T^d(Y^+)$.

2. *Precaución*: para gastos no planeados, debidos a la incertidumbre. *A priori* dependería de la renta y del tipo de interés, $M_P^d = M_P^d(Y^+, i^-)$, pero por simplicidad se asimila al motivo transacción.

3. *Especulación*: deriva de las expectativas sobre la evolución futura del tipo de interés (y, por tanto, del precio de los bonos) y se debe a la función del dinero como depósito de valor, $M_E^d = M_E^d(i^-)$.

La demanda por motivo especulación es la aportación novedosa de Keynes, pues los otros motivos ya los habían contemplado los neoclásicos, y es la que liga la demanda de dinero al tipo de interés. Para explicar por qué se mantiene riqueza en dinero, activo "estéril", y no en bonos, que rinden un interés, Keynes se apoya en dos premisas:

1. *Relación inversa*: entre la tasa de rentabilidad de un bono y su precio de mercado[23][24].

2. *Valor crítico*: cada individuo piensa que el tipo de interés tiende hacia un valor crítico o tasa normal esperada, i^e[25].

De aquí resulta una demanda especulativa individual discontinua[26] (figura A4.3, panel a):

- *Tipo alto* ($i > i^e$): el individuo espera que el tipo caiga y que el precio de los bonos suba, y mantiene toda su riqueza en bonos ($M_E^d/P = 0$).

- *Tipo bajo* ($i < i^e$): espera que el tipo suba y que el precio de los bonos caiga, y mantiene toda su riqueza en dinero ($M_E^d/P = W$).

Al agregar individuos con valores críticos dispares, la demanda especulativa agregada resulta una función continua, decreciente y convexa del tipo de interés.

Función de demanda de dinero: de la suma de los tres motivos se obtiene la demanda real de dinero agregada keynesiana, que en su versión lineal es:

$$\frac{M^d}{P} = L(Y^+, i^-) = kY - hi$$

donde $k > 0$ recoge la demanda por transacción y precaución como fracción de la renta real y $h > 0$ la sensibilidad de la demanda especulativa al tipo de interés, coste de oportunidad de mantener saldos monetarios. Keynes supone que h es elevada y que el componente especulativo, gobernado por expectativas volátiles, hace muy inestable la demanda de dinero, a diferencia del término kY. Lo esencial es que la demanda de dinero ya no solo depende de la renta, sino también del tipo de interés, entendido no como una recompensa a la abstinencia del consumo, sino como una retribución por renunciar a la liquidez[27].

El otro eslabón débil: la trampa de liquidez: es un estado de la economía en el cual los agentes estiman que el tipo de interés ha alcanzado un suelo y ha de subir. Cada agente cuyo valor crítico supere el tipo efectivo guarda toda su riqueza en dinero, y cuanto más baja el tipo, más agentes se encuentran en ese caso. Por ello, si el tipo es lo suficientemente bajo (figura A4.3, panel b), la demanda de dinero se vuelve perfectamente elástica al tipo de interés ($h \to \infty$). Cualquier aumento de la cantidad de dinero se mantiene entonces en efectivo, sin destinarse a bonos ni rebajar el tipo de interés, por lo que la política monetaria es totalmente inefectiva y la

[23] Hay una relación inversa y convexa entre el precio de mercado de un bono sin opción y su rentabilidad al vencimiento siempre y cuando el valor nominal, el cupón y el vencimiento no cambien (es decir, el bono es un bono emitido).

[24] Keynes consideró el caso más sencillo de bonos perpetuos: $P_B = \text{cup}/r_B$.

[25] Keynes subrayó que ese valor crítico o tasa "normal" es en gran medida convencional, fruto de la psicología y las expectativas del mercado (Keynes, 1936), y en la *Teoría general* renunció expresamente al tipo de interés natural wickselliano como referencia útil. Por simplicidad, se identifica aquí el tipo crítico con la tasa normal esperada. En rigor, como el cupón compensa parte de la pérdida de capital esperada, el tipo crítico es $i_c = i^e/(1 + i^e) < i^e$.

[26] ¿Qué ocurre cuando $i = i^e$? En este caso el individuo es indiferente entre ambos activos financieros y podría mantener cualquier combinación de ellos. Por tanto, la función de demanda (real) de dinero (individual) se volvería perfectamente elástica. No obstante, Keynes no contempla esta situación.

[27] "The rate of interest at any time, being the reward for parting with liquidity, is a measure of the unwillingness of those who possess money to part with their liquid control over it" (Keynes, 1936).

fiscal, especialmente potente. Sin embargo, Keynes presentó la trampa de liquidez como un caso límite del que no conocía ejemplos[28].

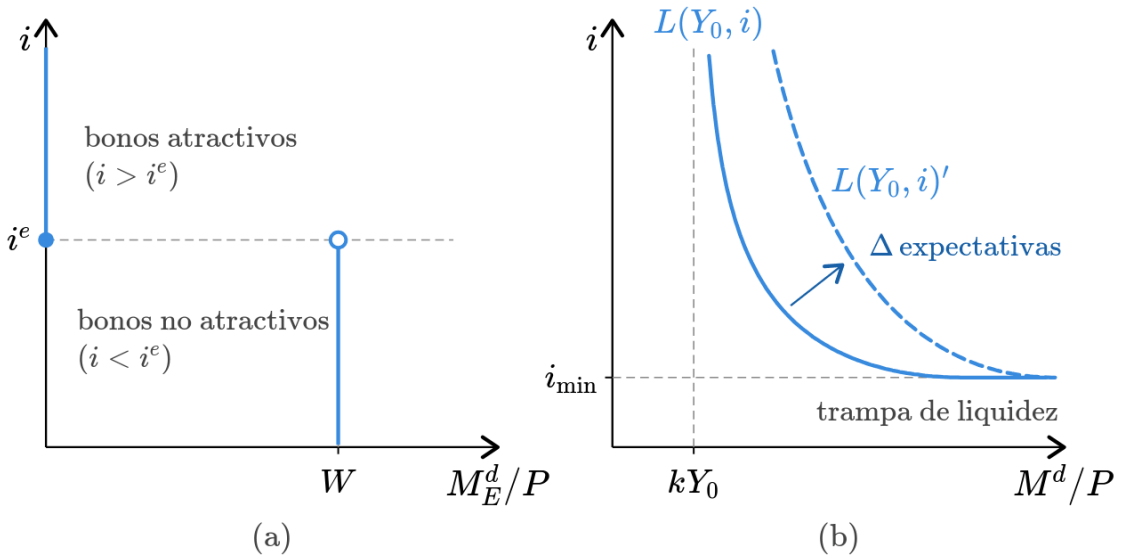

Figura A4.3: Demanda de dinero keynesiana. La demanda especulativa individual, M_E^d/P, es nula si $i > i^e$ y igual a la riqueza, W, si $i < i^e$, de modo que es discontinua en el tipo crítico i^e (a). La demanda agregada, $L(Y_0, i)$, tiende a kY_0 con tipos altos, es decreciente y convexa y se vuelve horizontal en $i_{\min}$, la trampa de liquidez (b). La curva discontinua, $L(Y_0, i)'$, es su desplazamiento ante un cambio de expectativas.

Curva LM: dado que la oferta real de dinero es exógena, M^s/P, la relación LM recoge los pares renta-tipo de interés que mantienen el mercado de dinero (y, por la ley de Walras, el de bonos) en equilibrio, $M^s/P = kY - hi$[29]:

$$i = \frac{1}{h}\left[kY - \frac{M^s}{P}\right]$$

Gráficamente, en el plano renta-tipo de interés, la LM es una recta creciente, de pendiente k/h, porque al aumentar la renta aumenta la demanda de dinero por motivo transacción y el tipo de interés sube para racionarla. Para Keynes, por tanto, el tipo de interés se determina por fuerzas monetarias, en el mercado de dinero, y no por fuerzas reales en el mercado de fondos prestables como para los neoclásicos. Cuanto mayor es h, más plana es la LM, que es horizontal en la trampa de liquidez. El debate con los monetaristas, para quienes $h \approx 0$ y la LM es casi vertical, se recoge en el tema A05, apartado 2.2.4.

[28] Keynes habla en el capítulo 15 de una preferencia por la liquidez que puede volverse "virtualmente absoluta", pero no usa la expresión *trampa de liquidez*, que se debe a Robertson (*Essays in Monetary Theory*, 1940).

[29] Una formulación posterior, ajena a Keynes, sustituye la LM por una regla de tipo de interés del banco central y deja que la cantidad de dinero, M^s, se ajuste (Romer, 2000).

La teoría de la preferencia por la liquidez (supuestos, modelo, desarrollo gráfico y valoración) se desarrolla en el tema A35, apartado 1.2, y la derivación de la relación LM, su estática comparativa y su análisis gráfico, en el tema A05, apartado 1.2.1.

2.3.4. Equilibrio en el mercado de bienes y dinero

Equilibrio: en el plano renta real-tipo de interés (figura A4.4), el cruce de las curvas IS y LM da el equilibrio de la economía por el lado de la demanda agregada.

Estática comparativa:

▪ *Política monetaria*: un aumento de la oferta monetaria desplaza la LM hacia la derecha, de LM_0 a LM_1, reduce el tipo de interés y eleva la renta, tanto más cuanto más inclinada sea la LM (menor h) y más plana la IS (mayor v).

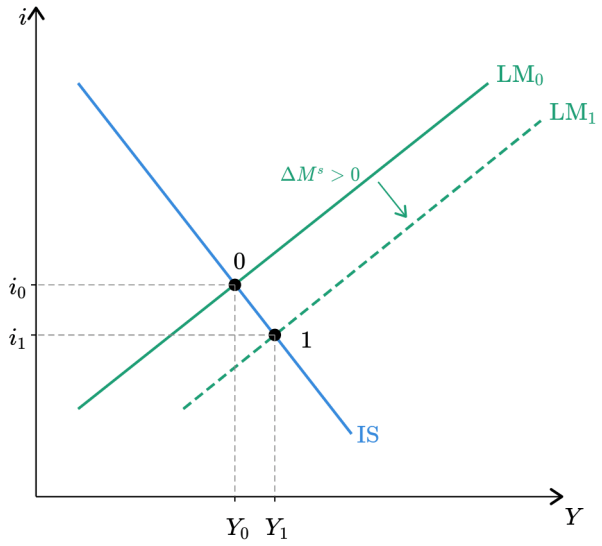

Figura A4.4: Equilibrio IS-LM y política monetaria expansiva. El cruce de la IS y la LM_0 da el equilibrio inicial, punto 0, con renta Y_0 y tipo de interés i_0. Un aumento de la oferta monetaria ($\Delta M^s > 0$) desplaza la LM a la derecha, hasta LM_1, y en el nuevo equilibrio, punto 1, el tipo de interés baja a i_1 y la renta sube a Y_1.

▪ *Política fiscal*: un aumento del gasto público desplaza la IS hacia la derecha y eleva la renta y el tipo de interés. Sin embargo, la subida del tipo de interés expulsa inversión privada, de modo que, salvo en la trampa de liquidez, el efecto sobre la renta es menor que el del multiplicador simple, $1/(1-c)$. El efecto sobre la renta es tanto mayor cuanto más plana sea la LM (mayor h) y más inclinada la IS (menor v).

2.3.5. Ajuste en el mercado keynesiano: la inversión no planeada y la preferencia por la liquidez

Ajuste en el mercado de bienes y servicios: el aspa keynesiana y la determinación de la renta: la proposición fundamental de Keynes es que los excesos de demanda en el mercado de bienes y servicios no los equilibran los precios, sino las cantidades. En esta interpretación, el consumo planeado siempre coincide con el efectivo y la variable de ajuste no es el tipo de interés real, como en los clásicos, sino la diferencia entre la inversión planeada y la efectiva. Los cambios de expectativas se materializan, por tanto, en la acumulación o desacumulación de existencias.

Gráficamente, la idea se ilustra en el plano renta-demanda agregada con la llamada *aspa keynesiana* (figura A4.5):

1. *Línea de 45°*: recoge los puntos en los que el gasto planeado equivale al gasto efectivo.

2. *Demanda agregada*: recoge el gasto planeado. Su pendiente es $\partial \mathrm{DA}/\partial Y = c$, y la ordenada en el origen es el gasto autónomo para un tipo de interés dado, $A_0 - vr$, con $A_0 = C_0 + I_0 + G_0$.

3. *Equilibrio*: en la intersección de ambas curvas, la inversión planeada coincide con la efectiva y la demanda agregada, con la renta. Este equilibrio es único y estable.

- *Equilibrio sin pleno empleo*: nada garantiza que el equilibrio se alcance en el nivel de renta de pleno empleo, Y_{PE}, lo que solo sucederá si la demanda agregada es suficiente.

- *Estabilidad*: fuera del equilibrio, el ajuste es vía existencias. A la derecha del equilibrio, el producto (gasto efectivo) es superior a la demanda agregada (gasto planeado), se acumulan existencias no deseadas en cuantía $Y_1 - \mathrm{DA}(Y_1) = (1-c)(Y_1 - Y^*)$ y las empresas reducen su producción hasta Y^*. A la izquierda ocurre lo contrario.

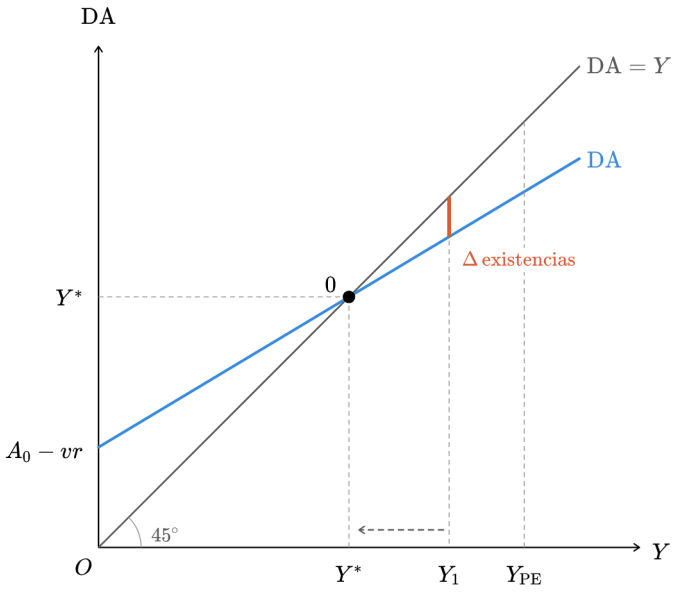

Figura A4.5: Aspa keynesiana. La renta de equilibrio, Y^*, iguala la demanda agregada y la producción. A su derecha se acumulan existencias no deseadas y la producción se reduce. Nada garantiza, sin embargo, que Y^* alcance la renta de pleno empleo, Y_{PE}.

Ajuste en el mercado de dinero: el tipo de interés: a diferencia del mercado de bienes y servicios, aquí el precio (el tipo de interés) sí desempeña un papel central, debido a la teoría de la preferencia por la liquidez. El tipo de interés actúa, con todo, en conjunción con las expectativas.

2.4. Macromodelo keynesiano OA-DA

El modelo se completa al levantar el supuesto de precios fijos. La oferta agregada, derivada del mercado de trabajo, y la demanda agregada, derivada del IS-LM, determinan entonces conjuntamente la producción y el nivel de precios.

2.4.1. Oferta y demanda agregada

Oferta agregada: como en el pensamiento neoclásico, se obtiene al sustituir el nivel de empleo de equilibrio en la función de producción agregada, aunque la relación causal es inversa (las empresas ajustan su producción a la demanda agregada). En este caso presenta tres tramos, debido a la forma de la curva de oferta de trabajo:

1. *Primer tramo*: los aumentos de precios elevan la demanda de trabajo y el empleo, pero no el salario nominal, pues la ilusión monetaria impide indexarlo a la inflación.

2. *Segundo tramo*: la ilusión monetaria "se diluye" y los aumentos de precios y de la demanda de trabajo sí elevan el salario nominal.

3. *Tercer tramo*: es vertical en el nivel de pleno empleo, como la oferta agregada neoclásica (apartado 1.5.1).

El desempleo involuntario surge de una demanda agregada insuficiente, con el equilibrio del mercado de trabajo en el primer tramo (§2.2.3). Para Keynes, por ello, la principal cuestión de la teoría económica no era la asignación de recursos, sino explicar por qué una economía puede tener persistentemente un ingreso inferior al de pleno empleo y cómo emplear los recursos ociosos. Según la lectura posterior de la síntesis, la OA acabaría siendo vertical a largo plazo en el plano (Y, P). Keynes, sin embargo, desconfiaba de los razonamientos que remiten al ajuste a largo plazo. En *A Tract on Monetary Reform* (1923), al discutir la teoría cuantitativa, ya había advertido que el largo plazo es una guía engañosa para los asuntos corrientes, porque "a largo plazo todos estaremos muertos".

Demanda agregada: se obtiene de la proyección en el plano renta real-precios de cada uno de los equilibrios en el mercado de bienes, servicios y dinero.

Equilibrio: el corte de la OA y la DA fija la producción y el nivel de precios, y puede darse en un punto en el que existen recursos ociosos.

Desarrollo gráfico: el macromodelo keynesiano se representa en el plano renta real-precios, con los paneles auxiliares del mercado de trabajo, la función de producción y la recta de 45° (figura A4.6):

1. *Mercado de trabajo*: en el plano empleo-salario nominal, la oferta de trabajo, L^s, es horizontal en el salario mínimo w_m, después creciente y vertical en el pleno empleo, $\bar{L}$. La demanda de trabajo, L^d, es decreciente.

2. *Función de producción*, $f(L)$: relaciona el empleo con la producción, de modo que producir Y^* requiere el empleo L^*, con $Y^* = f(L^*)$, y al pleno empleo le corresponde $\bar{Y} = f(\bar{L})$.

3. *Recta de 45°*: traslada la producción que fija el plano renta real-precios al eje vertical de la función de producción.

4. *Oferta agregada*, OA: creciente hasta $\bar{Y}$, porque una subida de precios abarata el salario real y eleva el empleo, y vertical a partir de ese nivel. No es plana ni siquiera frente al tramo horizontal de L^s, porque, con el salario nominal fijo en w_m, contratar más exige que la productividad marginal, decreciente, iguale un salario real menor, $w_m/P = f'(L)$, es decir, un precio mayor. A partir de L_1 sube además el salario nominal y la OA se empina.

5. *Demanda agregada*, DA: proyecta los equilibrios IS-LM para cada nivel de precios y es decreciente, porque un nivel de precios más alto reduce la oferta real de dinero y desplaza la LM hacia la izquierda.

6. *Equilibrio*: en el corte de la OA y la DA, con producción Y^* y precios P^*. En el mercado de trabajo, la demanda corta a la oferta en su tramo horizontal, con empleo $L^* < \bar{L}$, paro involuntario U_I y paro voluntario U_V. Las flechas verdes recogen el sentido de la causalidad, inverso al del macromodelo neoclásico. La demanda determina la producción, Y^*, la recta de 45° la lleva a la función de producción y esta da el empleo necesario para obtenerla, L^*, que es el que se observa en el mercado de trabajo.

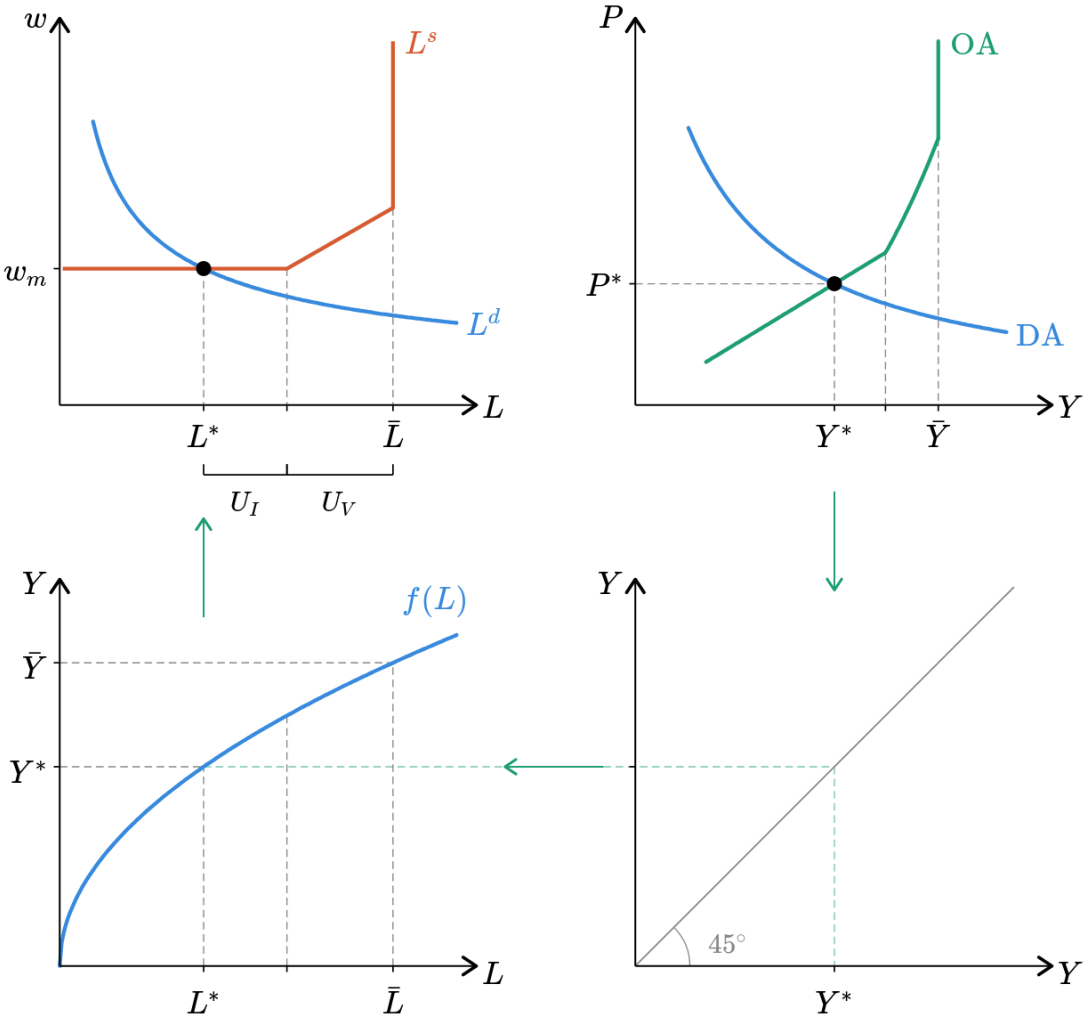

Figura A4.6: Macromodelo keynesiano OA-DA. Con el salario nominal rígido a la baja en w_m, la oferta agregada tiene pendiente positiva hasta el pleno empleo y, junto con la demanda agregada, derivada del IS-LM, determina la producción, Y^*, y el nivel de precios, P^*. En el equilibrio representado, la demanda de trabajo corta la oferta en su tramo horizontal, de modo que el empleo es $L^* < \bar{L}$, con paro involuntario U_I y voluntario U_V. El nivel de pleno empleo es $\bar{L}$ y la producción de pleno empleo, $\bar{Y} = f(\bar{L})$.

Estática comparativa:

▪ *Políticas de demanda*: una expansión fiscal o monetaria desplaza la DA hacia la derecha. Mientras existan recursos ociosos, eleva la producción, el empleo y los precios, tanto más la producción y menos los precios cuanto más plana sea la OA, y en el tramo vertical de pleno empleo solo eleva los precios. La magnitud del desplazamiento de la DA depende de las pendientes de la IS y la LM (§2.3.4).

▪ *Perturbación de oferta*: una subida del salario nominal mínimo, w_m, o una caída de la productividad del trabajo desplaza la OA hacia la izquierda, con lo que suben los precios y caen la producción y el empleo.

2.4.2. La política económica en Keynes

Para Keynes, evitar el paro involuntario exige un uso activo del presupuesto, que supla la demanda que falte, y unos tipos de interés bajos que animen la inversión productiva. Algunas de sus recomendaciones concretas aparecen en la *Teoría general*:

1. *Tipos de interés*: deben seguir bajos mientras el capital sea escaso, para que su acumulación continúe hasta que deje de serlo. Cuando el capital dejara de ser escaso llegaría la "eutanasia del rentista", es decir, el fin del poder del capitalista para obtener una renta de la mera escasez del capital.

2. *Consumo*: Keynes aboga por dirigir la propensión a consumir, en parte a través del sistema impositivo.

3. *Inversiones*: Keynes aboga por una "socialización bastante completa de la inversión", es decir, por que el Estado asuma un papel rector en el volumen de inversión. Con todo, rechaza explícitamente el socialismo de Estado, porque lo importante no es la propiedad pública de los medios de producción, sino el volumen total de la inversión.

2.4.3. La política económica de la síntesis

La síntesis neoclásica (SNC), dominante hasta su quiebra en los años setenta (§3.2), da la razón a Keynes en el corto plazo y a los neoclásicos en el largo:

1. *Corto plazo*: mientras precios y salarios no se ajustan, la política de demanda discrecional suple la falta de demanda agregada, y la fiscal es la más eficaz, por los dos eslabones débiles.

2. *Largo plazo*: el ajuste de precios y salarios devuelve la economía al pleno empleo. Basta entonces con un presupuesto reducido y en equilibrio, y con una política monetaria que provea liquidez y mantenga la inflación bajo control.

2.5. Conclusiones

Aspectos positivos: la aportación de Keynes transformó la teoría económica:

- *Desempleo involuntario*: explica por qué fue tan elevado en muchos países durante los años treinta.

- *Efectos reales de las variables monetarias*: explica cómo se producen, con lo que rompe con la dicotomía clásica.

- *Intervención pública*: ofrece una justificación fundamentada de la actuación del sector público en la economía.

- *SNC*: en la reinterpretación de la síntesis, de gran utilidad pedagógica por su simplicidad, Keynes explica el corto plazo, en el que la política económica es efectiva.

Limitaciones:

- *Metodológicas*: Keynes no formalizó muchas de sus ideas, lo que dio lugar a múltiples debates sobre "qué es lo que quería decir Keynes exactamente" (sección 3).

- *Núcleo teórico*: deja fuera los retardos con que actúa la política económica.

3. Reacciones y desarrollos posteriores

La obra de Keynes y su difusión por los primeros autores de la SNC situaron la macroeconomía en el centro de la teoría económica y suscitaron varias reacciones.

3.1. El debate sobre los saldos reales

El debate de los saldos reales versó sobre los efectos de las variaciones del nivel de precios (en particular, de la deflación) sobre la demanda agregada y la capacidad de la economía para volver por sí sola al pleno empleo.

3.1.1. La postura de Keynes: efecto Keynes (1936)

Origen: en la *Teoría general* (1936), Keynes examina si una bajada de los salarios nominales y de los precios puede devolver por sí sola la economía al pleno empleo, como sostenía la tesis neoclásica.

Idea: la deflación solo estimula la demanda de forma indirecta, a través del tipo de interés, como lo haría una expansión monetaria.

Razonamiento: si la cantidad nominal de dinero no cambia, la deflación eleva los saldos reales en manos del público. Con más dinero real, el tipo de interés real baja, la inversión sube y el multiplicador traslada ese impulso a la renta (*efecto Keynes*), lo que en el plano renta-tipo de interés equivale a desplazar la LM hacia la derecha (figura A4.7).

Como el canal de transmisión es el tipo de interés, la deflación puede no estimular la demanda en una trampa de liquidez o con una inversión insensible al tipo de interés, y tiene además efectos adversos sobre las deudas y las expectativas (§2.2.1). Por ello, Keynes defendía la estabilidad de los salarios nominales, prefería la expansión monetaria a la bajada de salarios y abogaba por estímulos fiscales en un contexto deflacionario.

Conclusiones:

- *Aspectos positivos*: muestra que la flexibilidad salarial no ofrece nada que la política monetaria no pueda lograr con menos costes.

- *Limitaciones*: su eficacia depende de las pendientes de la IS y de la LM, y deja fuera el efecto de la riqueza real sobre el gasto que señalaría Pigou (§3.1.2).

3.1.2. La postura neoclásica: efecto Pigou (1943) y efecto Metzler (1951)

Origen: Pigou (1943) consideraba que Keynes no tenía en cuenta el efecto de los saldos monetarios reales en el consumo corriente, que depende de la riqueza real y no solo de la renta (§1.3.1). La función de consumo es entonces $C = C(Y^+, r^-, W^+)$, donde $W = M^s/P + B/P$ es la "riqueza real neta", suma de la oferta monetaria y las tenencias de bonos, en términos reales.

Idea: la deflación revaloriza la riqueza real de los agentes, que pueden gastar más, de modo que un canal de riqueza lleva el impulso al consumo y a la demanda.

Razonamiento: con este efecto Pigou[30] o efecto riqueza, el tipo de interés pierde su condición de único puente entre el mercado financiero y el de bienes. Por ello, los dos eslabones débiles (la trampa de liquidez y la escasa sensibilidad de la inversión al tipo de interés) ya no impiden el retorno automático al pleno empleo. En términos gráficos (figura A4.7), el efecto Pigou supone un desplazamiento hacia la derecha de la curva IS, junto con el desplazamiento de la curva LM provocado por el efecto Keynes. Así, aunque en la trampa de liquidez o con una IS muy inclinada el efecto Keynes apenas eleve la renta (de 0 a 0′), el desplazamiento de la IS lleva la economía al pleno empleo, $\bar{Y}$ (punto 1).

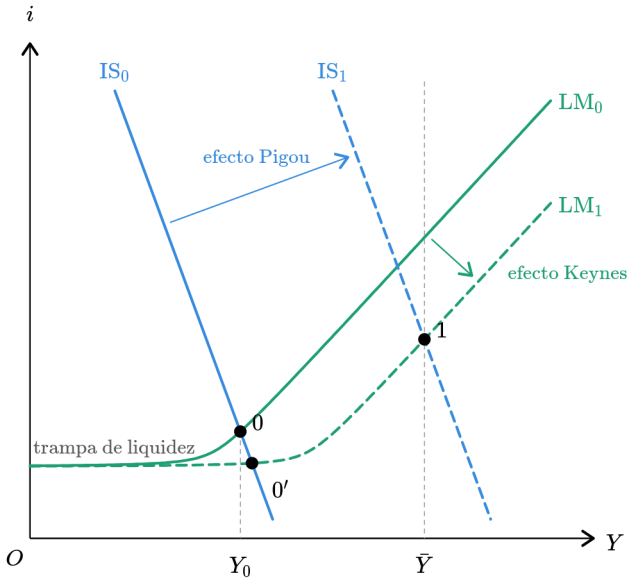

Figura A4.7: Efectos Keynes y Pigou en el IS-LM. Con la IS muy inclinada y la LM cerca de la trampa de liquidez, la deflación desplaza la LM de LM_0 a LM_1 (efecto Keynes) y la renta apenas varía (de 0 a 0′). El efecto Pigou desplaza además la IS de IS_0 a IS_1 y lleva la economía a la renta de pleno empleo, $\bar{Y}$ (punto 1).

El efecto riqueza hace a la economía más estable de lo que sugería Keynes. Para el paradigma neoclásico era a la vez una oportunidad, porque reforzaba el retorno automático al pleno empleo, y un reto, porque rompía la dicotomía clásica.

Extensiones:

▪ *Efecto Metzler*: Lloyd Metzler (1951) incluye en la riqueza real el *stock* de capital, K, en forma de títulos cuyo valor depende del tipo de interés. Como el ahorro depende de la riqueza, importa también su composición (*efecto Metzler*). Por ejemplo, una compra de títulos en mercado abierto por el banco central reduce la riqueza privada real, eleva el ahorro y rebaja permanentemente el tipo de interés real de equilibrio. La compra no es, por tanto, neutral, y el tipo de interés deja de depender solo de factores reales.

▪ *Controversia Patinkin*: Don Patinkin (1956) intenta reconciliar el efecto Pigou con la neutralidad monetaria. Demuestra que, para mantener ambos a largo plazo, es lógicamente necesario violar la dicotomía clásica, lo que dio lugar a la llamada *controversia Patinkin*.

[30] El efecto fue denominado así por Don Patinkin en un artículo de 1948. No confundir con el efecto riqueza directo.

▪ *Equivalencia ricardiana*: Robert Barro (1974) debilita el efecto Pigou con su revitalización de la equivalencia ricardiana, ya que, si los bonos públicos no constituyen riqueza neta de los individuos, el efecto riqueza queda reducido a los saldos monetarios reales.

Conclusiones:

▪ *Aspectos positivos*: al ligar el gasto a la riqueza, muestra que, en el plano teórico, el equilibrio con desempleo de Keynes exige alguna rigidez de precios o salarios. Además, el efecto Pigou deja a la síntesis neoclásica el canal riqueza que esta incorporó al modelo IS-LM.

▪ *Limitaciones*: la bajada de precios debe ser muy elevada para que el efecto actúe, y el análisis no tiene en cuenta los efectos de la deflación sobre las expectativas. Además, como objetó Kalecki (1944)[31], el efecto Pigou solo opera sobre el dinero externo, porque la deflación eleva la carga real de las deudas privadas y lo que ganan los acreedores lo pierden los deudores, con el riesgo de quiebras generalizadas.

3.2. La síntesis neoclásica

Keynes había roto la dicotomía clásica y abierto varios canales de transmisión de las perturbaciones por el lado de la demanda agregada, pero su microfundamentación era escasa y, en el caso del consumo, la evidencia de Kuznets contradecía la teoría a largo plazo.

> La síntesis neoclásica se desarrolla en el tema A05, apartados 1.1 a 1.3, y cada uno de sus componentes, en su tema propio. La renta relativa y el ciclo vital se estudian en el tema A33, apartados 1.2.2 y 2.2, y el acelerador y la q de Tobin, en el tema A34, apartados 1.3 y 2.2.2. Los modelos de Baumol-Tobin, de costes de ruptura y de carteras mixtas se encuentran en el tema A35, apartado 1.3. La curva de Phillips y la ley de Okun, por último, se estudian en el tema A28, apartado 1.2, y, como teoría de la inflación, en el tema A41, apartado 2.2.1.

Por ello, los seguidores de la SNC microfundamentaron cada componente de la demanda agregada (consumo, inversión y demanda de dinero) y ampliaron los canales de transmisión del IS-LM con un canal riqueza (del que el efecto Pigou es un caso) que desplaza la IS.

Por el lado de la oferta, la **curva de Phillips** (Phillips, 1958; Lipsey, 1960), interpretada por Samuelson y Solow (1960) como un **menú de política económica**, y la **ley de Okun** (1962) parecían permitir el ajuste fino (*fine-tuning*) de la economía. Sin embargo, la curva solo funcionaba en cualquier horizonte si los trabajadores padecían ilusión monetaria y no incorporaban sus expectativas de inflación[32], y la estanflación de los años setenta[33] la deshizo y dio paso a la crítica monetarista.

3.3. Poskeynesianismo

Origen: la escuela reúne a discípulos y colaboradores de Keynes en Cambridge (Joan Robinson, Nicholas Kaldor, Piero Sraffa y, después, Luigi Pasinetti), a Michał Kalecki, que llegó por su

[31] Kalecki, M. (1944), "Professor Pigou on 'The Classical Stationary State': A Comment", *The Economic Journal*, 54(213), pp. 131-132.

[32] Si el Gobierno expande la demanda para reducir el desempleo, las empresas ofrecen salarios nominales más altos ($g_w > 0$) y, con productividad constante, los trasladan a los precios ($g_w = g_P > 0$), de modo que el salario real no varía, aunque los trabajadores con ilusión monetaria perciben que sí.

[33] La estanflación, coexistencia de inflación y desempleo elevados, empezó a observarse en el Reino Unido a mediados de los sesenta y se generalizó con las crisis del petróleo de los setenta.

cuenta al principio de la demanda efectiva, y a una vertiente norteamericana (Sidney Weintraub, Paul Davidson, Hyman Minsky). Todos rechazan la lectura de la *Teoría general* que hizo la síntesis neoclásica, el "keynesianismo bastardo" de Robinson.

Idea: la demanda efectiva determina la producción y el empleo también a largo plazo, y la economía capitalista es inestable y opera en tiempo histórico irreversible, bajo incertidumbre no probabilizable. De ahí el rechazo de la síntesis, que reduce a Keynes a un caso particular de rigideces, y del equilibrio general walrasiano.

Razonamiento:

▪ *Vertiente británica*: pone el acento en la distribución, el crecimiento y el valor. Kaldor (1956) y Pasinetti (1962) hicieron depender el ahorro, y con él el crecimiento, del reparto entre salarios y beneficios (tema A43, apartado 1.2). Robinson y Sraffa sostuvieron frente a Samuelson y Solow la *controversia del capital*, sobre la coherencia de la función de producción agregada, y Sraffa (1960) formuló una teoría del valor neorricardiana.

Kalecki (1954) añadió los precios fijados con un margen sobre los costes, que depende del *grado de monopolio*, de modo que el reparto de la renta entre salarios y beneficios lo fija el poder de mercado de las empresas y no la productividad marginal. Su ecuación de beneficios, $\Pi = I + C_c$, donde C_c es el consumo de los capitalistas, se obtiene de la contabilidad nacional en una economía cerrada, sin sector público y en la que los trabajadores no ahorran. Resume la idea de que los capitalistas ganan lo que gastan, porque es la inversión, decidida de forma autónoma, la que determina los beneficios, y no al revés.

▪ *Vertiente norteamericana*: pone el acento en el dinero, la incertidumbre y las finanzas. Para Davidson, la irreversibilidad del tiempo y la incertidumbre hacen del dinero y de la preferencia por la liquidez el centro del análisis. Según la *hipótesis de inestabilidad financiera* de Minsky, en el auge la financiación pasa de posiciones cubiertas a especulativas y Ponzi, y la estabilidad genera su propia fragilidad (tema B26, apartado 1.2.1). Weintraub explicó la inflación por los costes salariales.

▪ *Dinero endógeno*: Kaldor (1970, 1982) y Moore (1988) sostienen que la cantidad de dinero no la fija el banco central, sino que la crea la banca al conceder préstamos, que generan a su vez depósitos. El banco central fija el tipo de interés y suministra las reservas que la banca le pide a ese tipo, de modo que la oferta monetaria es horizontal en el plano cantidad de dinero-tipo de interés (*horizontalismo*) y la cantidad la determina la demanda de crédito. Se invierte así la causalidad de la teoría cuantitativa, porque es la renta nominal la que arrastra al dinero, y pierde fundamento la LM con oferta exógena, y con ella el multiplicador monetario del tema A36, apartado 1.4, como explicación del proceso. Que los bancos centrales instrumenten hoy su política con el tipo de interés ha dado peso a esta descripción (McLeay, Radia y Thomas, 2014).

▪ *Un autor próximo*: Abba Lerner, keynesiano más que poskeynesiano, formuló las *finanzas funcionales* (1943), que juzgan la política fiscal por sus efectos sobre el empleo y la inflación, no por el saldo presupuestario.

Conclusiones:

▪ *Aspectos positivos*: su atención a la incertidumbre, a la fragilidad financiera y a la distribución recupera aspectos de la *Teoría general* que la síntesis dejó de lado, y la escuela explica de forma endógena la inestabilidad del capitalismo. Su descripción del dinero endógeno se acerca, además, a la forma en que hoy se instrumenta la política monetaria.

▪ *Limitaciones*: la principal es la ausencia de un modelo canónico compartido, que explica su escasa influencia en la corriente principal, pese al interés que despertó la hipótesis de Minsky tras la crisis *subprime* (2007-2009).

3.4. Keynesianos del desequilibrio

Junto al monetarismo, el propio keynesianismo produjo en los años sesenta y setenta una segunda crítica a la SNC, la de los **neokeynesianos del desequilibrio**. Sus figuras principales fueron Robert Clower y Axel Leijonhufvud, y la escuela rompió con el marco IS-LM para dar a Keynes otros cimientos microeconómicos. A su juicio, la teoría renta-gasto nada tiene que ver con lo que Keynes escribió, y comprenderlo exige dejar a Walras y regresar a Marshall.

3.4.1. Rechazo de la ley de Walras en un marco keynesiano (Clower, 1965)

Origen: Robert Clower (1965) abrió la escuela al rechazar que la ley de Walras rija para las demandas efectivas y microfundamentar así a Keynes.

Idea: Clower defiende que los modelos de equilibrio general walrasiano son el caso particular, sin racionamiento (demandas efectivas iguales a las nocionales), de modelos de equilibrio general no walrasianos que admiten transacciones a precios de no vaciado.

Razonamiento:

▪ *Demandas efectivas y nocionales*: la ley de Walras, identidad que se cumple para las demandas nocionales a cualquier vector de precios, solo se extiende a las demandas efectivas si existe un subastador walrasiano que coordina todos los planes. El subastador garantiza que los intercambios se realicen únicamente a precios de vaciado. Para Clower, Keynes pensaba en términos de decisiones efectivas. Si se admiten transacciones a precios de no vaciado, las decisiones *planeadas* o *nocionales*, en las que cada agente supone que toda su oferta (demanda) será atendida, no tienen por qué coincidir con las efectivas.

▪ *Hipótesis de decisión dual*: el subastador walrasiano implica, además, simultaneidad en las decisiones de oferta y demanda de un mismo agente y entre distintos agentes, el supuesto de la teoría walrasiana que, por contraste con la propuesta de Clower, suele denominarse *hipótesis de decisión unitaria*. Para microfundamentar el ajuste hasta el equilibrio no walrasiano, Clower la sustituye por la hipótesis de decisión dual, según la cual las decisiones son secuenciales, en dos etapas:

1. *Decisiones nocionales*: primero, los agentes toman sus decisiones walrasianas de oferta y demanda de bienes y trabajo, que a precios de no vaciado revelan los lados cortos y largos de los excesos.

2. *Decisiones efectivas*: después, deciden teniendo en cuenta no solo las restricciones presupuestaria y tecnológica, sino también las restricciones de cantidades.

Así, lo que cada agente llega a intercambiar depende de la renta que efectivamente ha obtenido, no solo de los precios. Los intercambios efectivos son, además, las únicas señales fiables que emiten los mercados.

Equilibrio: con rigideces de precios o en un horizonte de corto plazo, se pueden producir equilibrios con racionamiento o equilibrios no walrasianos, que formalizaron después Barro y Grossman (1971) y Malinvaud (1977), con tres características principales:

1. *Principio de intercambio voluntario*: no cabe obligar a ningún agente a demandar u ofrecer más de lo que desea.

2. *Fallos de coordinación*: las decisiones de los agentes son incompatibles entre sí, porque algunos agentes están racionados en algún mercado y no pueden llevar a cabo sus decisiones.

3. *Dominancia del lado corto*: formalmente, $x_i^* = \min\{x_i^d, x_i^s\}$, es decir, el "lado corto" de los excesos determina los intercambios efectivos y el "lado largo", el grado de racionamiento.

Corolario: los *spillovers* entre bienes y trabajo hacen que el desempleo involuntario no se deba solo al mercado de trabajo, y los equilibrios son persistentes y subóptimos.

Desarrollo gráfico: la hipótesis de decisión dual se representa en el plano trabajo-consumo (figura A4.8), partiendo de una situación de desequilibrio walrasiano:

1. *Función de producción*, $f(L)$: es la restricción tecnológica de las empresas.

2. *Restricción presupuestaria*: recoge las combinaciones de consumo y trabajo al alcance de los hogares, con pendiente w/P.

3. *Curvas de indiferencia*, u y u': recogen las preferencias de los hogares entre consumo y trabajo.

4. *Equilibrio*: difiere según exista o no el subastador walrasiano.

- *Mundo walrasiano* (punto 0): con precios de no vaciado (salarios nominales elevados y precios bajos), la decisión de consumo-trabajo de los hogares, que solo atiende a su restricción presupuestaria, no coincide con la de producción de las empresas, que solo atiende a la tecnológica. Hay un exceso de oferta de trabajo, de L^d a L^s, y otro de demanda de bienes, de Y^s a Y^d, que por la ley de Walras se compensan en valor y que el subastador elimina.

- *Mundo no walrasiano* (punto 1): con decisión dual, el lado corto de los excesos (la demanda de trabajo efectiva y la oferta de bienes efectiva) determina un equilibrio con racionamiento. En él, los hogares, racionados en el mercado de trabajo, ajustan su demanda efectiva de bienes a la renta realmente obtenida, de modo que solo el mercado de trabajo queda en desequilibrio. Este equilibrio es estable porque falta el subastador y porque, en una economía monetaria, el trabajo que no se vende no aporta renta con la que comprar bienes (Clower, 1967). El paro involuntario convive así con agentes que optimizan.

5. *Mercado de trabajo* (panel inferior): al salario real dado, la demanda de trabajo de las empresas, L^d, queda por debajo de la oferta de los hogares, L^s, y la diferencia es el paro. El salario real que vaciaría el mercado, $(w/P)_{\mathrm{W}}$, en el corte W de ambas curvas, es menor.

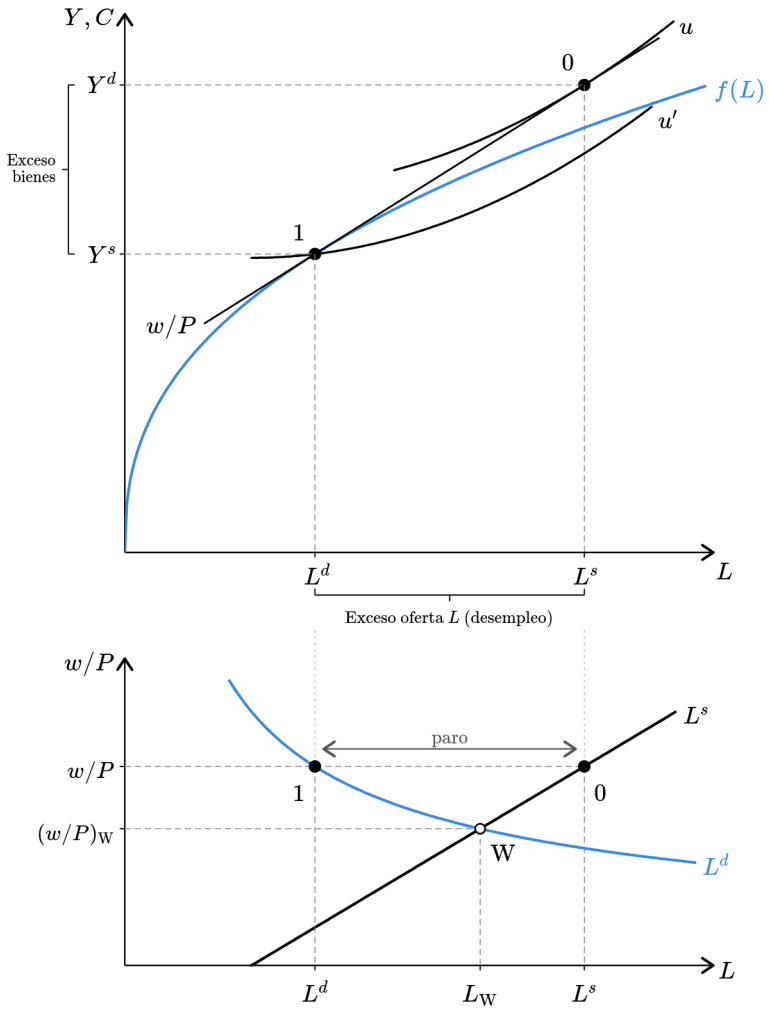

Figura A4.8: Hipótesis de la decisión dual (Clower, 1965). La demanda efectiva de bienes depende de la renta realmente obtenida en el mercado de trabajo, no de la renta nocional walrasiana. En el panel inferior, al salario real dado, la demanda de trabajo, L^d, queda por debajo de la oferta, L^s, y la diferencia es el paro.

Conclusiones:

▪ *Aspectos positivos*: microfundamenta la dependencia del consumo respecto de la renta efectivamente obtenida y, con ella, el multiplicador, y muestra que el desempleo involuntario es compatible con agentes racionales, sin ilusión monetaria.

▪ *Limitaciones*: toma como dados los precios de no vaciado sin explicar por qué no se ajustan, y deja sin fundamentar la rigidez que sostiene el resultado.

3.4.2. Amplificación de los desequilibrios: la retroalimentación de las recesiones (Leijonhufvud, 1967; 1968)

Origen: a raíz de las aportaciones de Clower, Axel Leijonhufvud (*On Keynesian Economics and the Economics of Keynes*, 1968) distingue la economía de Keynes de la "economía keynesiana" de la síntesis. Además, profundiza en el estudio del proceso de ajuste dinámico en un contexto de rigideces.

Idea: la aportación esencial de Keynes es haber invertido el orden habitual de las velocidades de ajuste, de modo que, sin subastador y con información imperfecta, las cantidades se ajustan antes que los precios. Esa lentitud del ajuste de los precios da lugar a un *efecto retroalimentación* que amplifica las perturbaciones.

Razonamiento: las transacciones a precios de no vaciado y la hipótesis de decisión dual permiten microfundamentar cómo una perturbación negativa de demanda tiene, partiendo del pleno empleo, un efecto contractivo amplificado sobre la producción y el empleo. Sirva de ejemplo una caída autónoma de la inversión debida al pesimismo empresarial:

- *Contexto walrasiano*: la caída de la inversión solo genera un desempleo friccional, igual a la diferencia entre la oferta y la demanda de trabajo walrasianas, que el subastador elimina de forma instantánea.

- *Contexto no walrasiano*: si precios, salarios y tipos de interés son rígidos y la información es imperfecta, el ajuste recae en las cantidades y se amplifica. Al no poder vender toda su producción, las empresas contratan menos trabajo. Los hogares que no encuentran empleo ingresan menos y recortan su demanda efectiva de bienes, lo que vuelve a restringir las ventas. El proceso se repite hasta un equilibrio no walrasiano con un paro mayor que el del ajuste walrasiano, y persistente por la lentitud del ajuste.

Corolario: los *shocks* de demanda se agrandan también por la vía financiera. La caída de la inversión genera incertidumbre, de modo que los agentes prefieren más liquidez y exigen más rentabilidad a los activos financieros (en particular, sube el tipo de interés a largo plazo), justo lo contrario de lo que haría falta para reactivar la economía.

Leijonhufvud defendía así el multiplicador keynesiano y la efectividad de las políticas de demanda frente a los equilibrios con racionamiento, cuyas posibilidades de política económica estudió después Edmond Malinvaud.

Conclusiones:

- *Aspectos positivos*: desplaza la explicación del desempleo keynesiano de la rigidez de los salarios a los fallos de información y de coordinación de una economía sin subastador, y separa la economía de Keynes de su lectura IS-LM.

- *Limitaciones*: es una reinterpretación eminentemente verbal, que no llegó a un modelo formal manejable, y su lectura de la *Teoría general* fue discutida por otros intérpretes de Keynes.

3.4.3. Política económica en el mercado de trabajo (Malinvaud, 1977)

Origen: Edmond Malinvaud (1977) se pregunta si una subida del salario real crea o destruye empleo, pues la subida del salario real actúa por dos lados:

- *Efecto oferta*: eleva los costes marginales de las empresas y puede desincentivar el empleo.

- *Efecto demanda*: puede estimular la demanda y, con ella, el empleo.

Idea: los efectos sobre el empleo de un cambio del salario real dependen de la naturaleza del paro. Malinvaud estudia tres equilibrios no walrasianos, el paro clásico, el paro keynesiano y la inflación reprimida, y trata el subconsumo como un cuarto caso límite.

Razonamiento:

- *Paro clásico*: el primer equilibrio no walrasiano se caracteriza por:
  - *Precios*: un salario real superior al salario real de vaciado ($(w/P)_{\mathrm{NW}} > (w/P)_{\mathrm{W}}$).
  - *Cantidades*: un exceso de demanda de bienes y un exceso de oferta de trabajo.
  - *Causa*: p. ej. un sindicato con poder negociador que fija un salario por encima del de vaciado. El lado corto del mercado de bienes es la oferta y el del mercado de trabajo es la demanda, de modo que tanto los consumidores (en el mercado de bienes) como los trabajadores (en el de trabajo) están racionados.
  - *Solución*: una bajada del salario real eleva el empleo, al aumentar la demanda de trabajo y la oferta de bienes de las empresas, mientras que un estímulo de demanda es ineficaz en este régimen.
- *Paro keynesiano*: se caracteriza por:
  - *Precios*: el salario real puede ser inferior, igual o superior al salario real de vaciado.
  - *Cantidades*: un exceso de oferta de bienes y un exceso de oferta de trabajo.
  - *Causa*: típicamente, una insuficiencia de demanda de bienes en una fase depresiva del ciclo. El lado corto de ambos mercados es la demanda, por lo que tanto las empresas (en el mercado de bienes) como los trabajadores (en el de trabajo) están racionados.
  - *Solución*: una subida limitada del salario real puede elevar el empleo, al estimular la demanda, y también lo hace la eliminación de la restricción de demanda efectiva (p. ej. impulso fiscal).
- *Otros resultados*: la tipología se completa con la *inflación reprimida* (excesos de demanda en ambos mercados, típicamente por controles de precios y salarios) y el *subconsumo* (exceso de demanda de trabajo y de oferta de bienes), un caso infrecuente.

Desarrollo gráfico: los equilibrios no walrasianos de Malinvaud se ilustran en el plano trabajo-producción de una economía de tipo Robinson Crusoe (figura A4.9), que permite representar ambos tipos de paro:

1. *Función de producción*, $f(L)$: recoge la tecnología de la empresa y es cóncava por la productividad marginal decreciente del trabajo.

2. *Curvas de indiferencia*, u_{W} y u_1: los hogares eligen entre consumo y trabajo según unas curvas convexas y crecientes, porque el trabajo es un "mal".

3. *Equilibrio*: el equilibrio walrasiano W se da en la tangencia entre ambas. En él, el salario real de vaciado, $(w/P)_{\mathrm{W}}$, iguala la productividad marginal del trabajo y la relación marginal de sustitución entre consumo y trabajo.

- *Paro keynesiano* (punto 1): la restricción de demanda efectiva limita las ventas a $Y_1 < Y_{\mathrm{W}}$ y el empleo a $L_1 < L_{\mathrm{W}}$. Los hogares, racionados en el mercado de trabajo, quedan en una curva de indiferencia más baja, u_1.

- *Paro clásico* (punto 2): con un salario real $(w/P)_{\text{NW}}$ superior al de vaciado, la empresa produce donde $f'(L_2) = (w/P)_{\text{NW}}$, con $L_2 < L_{\text{W}}$, y los hogares están racionados en ambos mercados.

4. *Mercado de trabajo* (panel inferior): la demanda nocional de trabajo, L^d, iguala la productividad marginal al salario real, y la oferta, L^s, la relación marginal de sustitución, y ambas se cortan en W. Con paro keynesiano, al salario real de vaciado la empresa solo contrata L_1, el empleo que permite la demanda efectiva, y el paro es $L_{\text{W}} - L_1$. Con paro clásico, al salario $(w/P)_{\text{NW}}$ la demanda cae a L_2 y la oferta sube a L^s_{NW}.

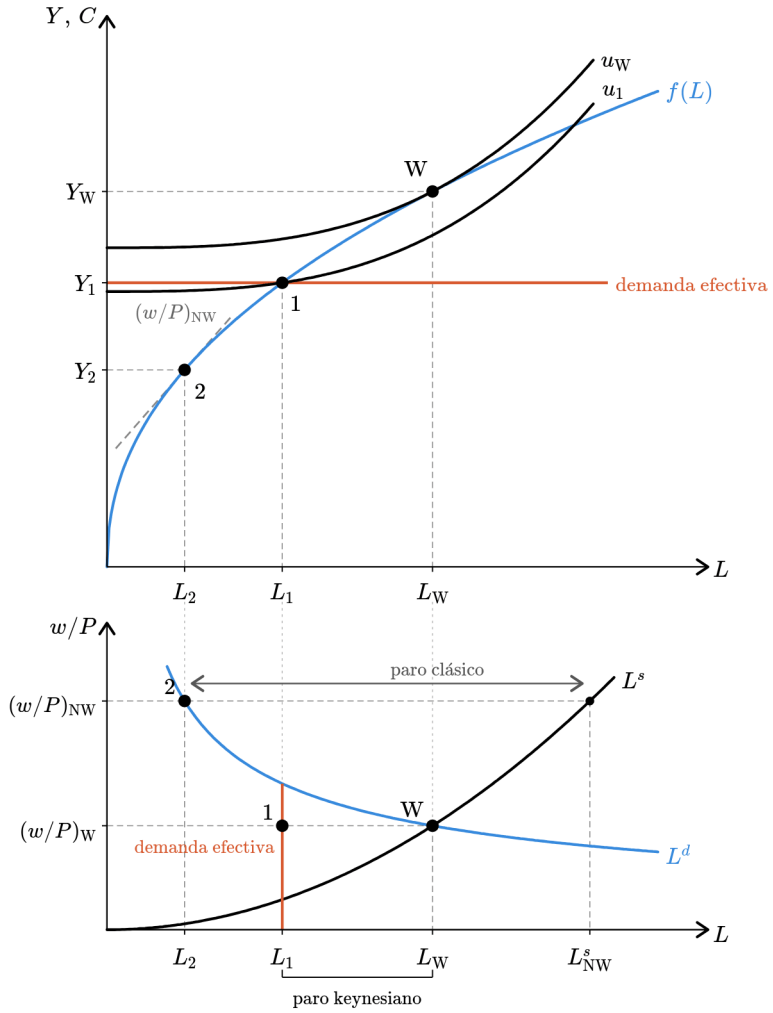

Figura A4.9: Paro keynesiano y paro clásico en una economía de Robinson Crusoe, en el plano trabajo-producción. El equilibrio walrasiano W se da en la tangencia entre la función de producción $f(L)$ y la curva de indiferencia u_{W}. En el punto 1 (paro keynesiano), la restricción de demanda efectiva limita la producción a Y_1 y el empleo a $L_1 < L_{\text{W}}$. En el punto 2 (paro clásico), un salario real $(w/P)_{\text{NW}}$ superior al de vaciado lleva a la empresa a producir donde $f'(L_2) = (w/P)_{\text{NW}}$, con $L_2 < L_{\text{W}}$. En el panel inferior se representa el mercado de trabajo, con el paro keynesiano, $L_{\text{W}} - L_1$, y el clásico, $L^s_{\text{NW}} - L_2$.

La tipología de Malinvaud, con la causa y la solución de cada régimen, se recoge también en el marco de la teoría del desempleo en el tema A28, apartado 1.4.

Conclusiones:

▪ *Aspectos positivos*: distinguir los regímenes de paro ayuda a diseñar la política económica, porque muestra que no hay políticas válidas para todos los casos (*one-size-fits-all*).

▪ *Limitaciones*: la prescripción depende de identificar correctamente la naturaleza del paro, es decir, qué lado es el corto en cada mercado, algo difícil de conocer cuando los precios de desequilibrio no son señales fiables. Además, la tipología es estática y no explica cómo se pasa de un régimen a otro.

3.4.4. Conclusiones

Aspectos positivos: la escuela analizó, a partir de una microfundamentación de Keynes, los mercados que no se vacían. De ese análisis salieron herramientas útiles para estudiar las rigideces, como la decisión dual o los fallos de coordinación. Explicar desde la microeconomía la rigidez de precios fue después el eje del programa de la nueva economía keynesiana, que tomó así el relevo.

Limitaciones: la escuela también recibió objeciones, de método y de fondo:

▪ *Metodología*: comparado con el IS-LM, su aparato gráfico pierde elegancia y alcance analítico. Además, la escuela desatiende la relación entre beneficios e inversión.

▪ *Núcleo teórico*: en cuanto a la reinterpretación de Keynes, cabe distinguir cuatro objeciones:

- *Oposición a Walras*: el blanco de la *Teoría general* eran autores de su tiempo, como Arthur Pigou, que dejaban sin explicar la formación de los precios. Keynes no dirige expresamente su crítica contra Walras y su subastador.
- *Rigidez a la baja*: a corto plazo, Keynes tampoco da por fijos los precios. Los considera flexibles y la rigidez a la baja la reserva para los salarios nominales (§2.1).
- *Duración de la rigidez*: sin saber cuánto duran los precios rígidos no hay base para intervenir, y la escuela no lo aclara.
- *Política económica*: las objeciones de los nuevos clásicos (inconsistencia dinámica, equivalencia ricardiana) alcanzaban a buena parte de sus recetas, por su carácter intervencionista.

Conclusiones

El tema contrapone dos maneras de entender el funcionamiento agregado de la economía. El macromodelo neoclásico descansa en dos líneas de defensa del pleno empleo, la flexibilidad del salario real, que vacía el mercado de trabajo, y la del tipo de interés real, que garantiza la ley de Say. Descansa también en la dicotomía clásica, que hace del dinero un velo. Sin embargo, la Gran Depresión mostró los límites de esos supuestos. Frente a ellos, Keynes invirtió la causalidad entre oferta y demanda, de modo que la demanda efectiva determina la producción y el empleo. A su vez, la inversión, gobernada por expectativas inciertas, determina el ahorro a través del multiplicador, y el tipo de interés se fija en el mercado de dinero como precio de la liquidez. De ahí la posibilidad de un equilibrio con desempleo involuntario que el mercado no corrige por sí solo.

Las reacciones a la *Teoría general* fijaron buena parte de la agenda de la macroeconomía. El debate sobre los saldos reales mostró que el equilibrio con desempleo exige alguna rigidez, y la síntesis neoclásica hizo de Keynes el caso de corto plazo de un modelo neoclásico. El monetarismo criticó después esa síntesis (tema A05). Los keynesianos del desequilibrio buscaron otra microfundamentación, en mercados que no se vacían, y su tipología de regímenes de desempleo reaparece en la teoría del mercado de trabajo (tema A28). Más tarde, la nueva economía keynesiana microfundamentó las rigideces nominales y reales (tema A07). Por su parte, las funciones keynesianas de consumo, inversión y demanda de dinero son el punto de partida de las teorías modernas de cada componente de la demanda (temas A33, A34 y A35).

La crisis financiera de 2008 devolvió actualidad a varias ideas de la *Teoría general*. Krugman (1998) y Eggertsson y Krugman (2012) reinterpretaron la trampa de liquidez como el límite inferior de los tipos de interés (*zero lower bound*), en el que la política monetaria convencional pierde eficacia y la fiscal la gana. Blanchard y Leigh (2013) concluyeron que los multiplicadores fiscales de la consolidación europea posterior a la crisis fueron mayores que los supuestos en las previsiones, y Summers (2014) recuperó el estancamiento secular de Hansen ante un tipo de interés natural persistentemente bajo. Por su parte, los modelos con hogares heterogéneos (Kaplan, Moll y Violante, 2018) devuelven a la propensión marginal a consumir un papel central en la transmisión de la política económica.

Anexo. Programas del macromodelo neoclásico

A.1. Oferta de trabajo

Planteamiento: formalmente, el problema de elección del consumidor-trabajador (§1.2.1) consiste en alcanzar la máxima utilidad posible dada su restricción de tiempo:

$$\max_{o,x} \; u = u(o,x) \quad \text{s.a.: } Px \leq w(l_0 - o) + M_0$$

donde x denota las unidades obtenidas de consumo, P su precio, w el salario nominal (precio del ocio), M_0 la renta monetaria exógena del agente y $l_0 = l + o$[34] refleja el reparto de la dotación de tiempo, l_0, entre la cantidad de horas destinada a trabajar, l, y la cantidad de horas destinada al ocio, o. El lagrangiano del problema es:

$$\mathcal{L} = u(o,x) + \lambda \left[w(l_0 - o) + M_0 - Px \right]$$

CPO: con solución interior y la restricción agotada, por el método de Lagrange:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \lambda} &= w(l_0 - o) + M_0 - Px = 0 \quad &\Rightarrow \; Px &= w(l_0 - o) + M_0 \\
\frac{\partial \mathcal{L}}{\partial x} &= \frac{\partial u}{\partial x} - \lambda P = 0 \quad &\Rightarrow \; \frac{\partial u}{\partial x} &= \lambda P \\
\frac{\partial \mathcal{L}}{\partial o} &= \frac{\partial u}{\partial o} - \lambda w = 0 \quad &\Rightarrow \; \frac{\partial u}{\partial o} &= \lambda w \\
& & \Rightarrow \; \text{RMS}_{xo} \equiv \frac{\partial u / \partial o}{\partial u / \partial x} &= \frac{w}{P}
\end{aligned}$$

Por tanto, en el óptimo el consumidor-trabajador gasta toda su renta e iguala su relación marginal de sustitución (RMS) consumo-ocio al salario real.

CSO: la tangencia es un máximo si las curvas de indiferencia consumo-ocio son estrictamente convexas (utilidad estrictamente cuasicóncava), pues la restricción presupuestaria es lineal. Bajo ese supuesto, la CPO es necesaria y suficiente.

Solución: a partir de la CPO se obtiene la función de oferta de trabajo individual, que, si el efecto sustitución domina al efecto renta, depende positivamente del salario nominal y negativamente del índice de precios de la economía:

$$l^s = l^s(w^+, P^-)$$

Por agregación, en sentido amplio, se obtiene la oferta de trabajo agregada, $L^s = \bigodot_{i=1}^{n} l_i^s$, creciente con el salario real (§1.2.1). El operador $\bigodot$ indica que se agregan las funciones individuales sin explicitar el mecanismo, generalmente una suma.

[34] Se puede expresar también en términos de la renta potencial: $Px + wo \leq wl_0 + M_0$.

A.2. Demanda de trabajo

Planteamiento: el problema de las empresas competitivas consiste en maximizar su beneficio sujetas a sus restricciones tecnológicas y económicas. Se considera el corto plazo:

$$\max_{L} \ \Pi = Pf(\bar{K}, L) - wL - r_K\bar{K}$$

donde Π es el beneficio de la empresa, L la cantidad de trabajo contratada, $\bar{K}$ el *stock* de capital, fijo en el corto plazo, $f(\bar{K}, L)$ la función de producción, P el precio del producto, w el salario nominal y r_K el precio del capital.

CPO:

$$\frac{\partial \Pi}{\partial L} = P\frac{\partial f}{\partial L} - w = 0 \ \Rightarrow \ \underbrace{\frac{\partial f}{\partial L}}_{\text{Pmg}_L} = \underbrace{\frac{w}{P}}_{\text{salario real}}$$

Es decir, cada empresa contratará trabajadores hasta que el valor de su producto marginal, $P\partial f/\partial L$, se iguale a su salario nominal o, dicho de otro modo, hasta que su productividad marginal, $\partial f/\partial L$, sea igual a su salario real, w/P.

CSO: la condición es suficiente porque, con rendimientos marginales decrecientes del trabajo ($f''_{LL} < 0$), el beneficio es estrictamente cóncavo en L y el punto crítico es un máximo.

Solución: de la CPO se obtiene la función de demanda de trabajo individual:

$$l^d = l^d(w^-, P^+)$$

Por agregación se obtiene la demanda de trabajo agregada, $L^d = \bigodot_{i=1}^{n} l_i^d$, decreciente con el salario real (§1.2.2).

A5. La síntesis neoclásica y el monetarismo

La síntesis neoclásica. El monetarismo.

Introducción

Introducción común. Los temas A02 a A07 comparten una introducción, que se desarrolla al comienzo del tema **A02**. Recorre la evolución del pensamiento económico, del mercantilismo y la fisiocracia a la escuela clásica, la revolución marginalista, Keynes y la síntesis neoclásica, el monetarismo, la nueva macroeconomía clásica, la nueva economía keynesiana y la nueva síntesis neoclásica.

"Los objetivos básicos a largo plazo, compartidos, estoy seguro, por la mayoría de los economistas, son la libertad política, la eficiencia económica y una igualdad sustancial del poder económico [...]. Creo que los tres objetivos pueden alcanzarse mejor si se confía, en la medida de lo posible, en un mecanismo de mercado dentro de un 'orden competitivo' para organizar la utilización de los recursos económicos".

Con estas palabras, Milton Friedman, en *A Monetary and Fiscal Framework for Economic Stability* (1948), defendía su postura de política económica frente al intervencionismo que había impulsado el keynesianismo de posguerra. Sin embargo, hubo que esperar al menos dos décadas para que el monetarismo, término acuñado por Karl Brunner en 1968, se materializase en recetas de política económica aplicables a los problemas de la economía real. Su postulado distintivo no era la confianza en los mercados, sino la estabilidad de la demanda de dinero y, como contrapartida, la inestabilidad que la política discrecional inducía en una economía privada que, dejada a sí misma, tendía al equilibrio.

A partir de mediados de los cincuenta, un círculo de economistas con base principal en la Universidad de Chicago y marcado por la figura de Friedman (premio Nobel en 1976) combatió con firmeza muchas tesis keynesianas. Su doctrina, el monetarismo[1], conserva buena parte del legado neoclásico y solo rompe con él en el papel del dinero, neutral a largo plazo pero no a corto (apartado 2.1).

El núcleo del **monetarismo** es la teoría cuantitativa del dinero, que Friedman (1956) reformuló como una teoría de la demanda de dinero estable en la tradición de los saldos de caja de Cambridge, si bien la escuela recoge también la herencia de las teorías monetarias del ciclo anteriores, como las de Wicksell o Keynes. Los monetaristas combaten la síntesis neoclásica en

[1] Una caracterización habitual define el monetarismo como la corriente de pensamiento convencida de que la demanda agregada, medida en términos nominales, depende ante todo del dinero y, más en concreto, del ritmo al que varía su cantidad.

el terreno de esta y con su mismo instrumental[2] (IS-LM, curva de Phillips, demanda de dinero y demanda agregada), pese a presentarse como continuadores de los neoclásicos.

En paralelo a esta disputa, y en gran medida ajena a ella, discurrió la controversia del capital entre las dos Cambridge (Joan Robinson y Piero Sraffa frente a Paul Samuelson y Robert Solow), un debate sobre los fundamentos de la función de producción agregada y de la teoría marginal de la distribución (tema A04, apartado 3.3).

El tema se organiza en tres partes:

- *Síntesis neoclásica*: sus principales rasgos y su análisis del lado de la demanda y del lado de la oferta.

- *Monetarismo*: sus principales rasgos y su crítica a la síntesis por el lado de la demanda y por el de la oferta.

- *Desarrollos posteriores*: la nueva macroeconomía clásica y las dos generaciones de la nueva economía keynesiana, que se estudian en sus temas propios.

[2] El rechazo de la mayoría de las herramientas keynesianas y de la síntesis neoclásica llegará después, con la nueva macroeconomía clásica, heredera del monetarismo.

1. La síntesis neoclásica

La exposición de la síntesis neoclásica recorre sus principales rasgos, el análisis del lado de la demanda y el de la oferta.

1.1. Principales rasgos

La **síntesis neoclásica** (SNC) es la corriente que, entre los años cuarenta y sesenta, se convirtió en la macroeconomía dominante al reconciliar la *Teoría general* de Keynes (1936) con la tradición neoclásica. La denominación se debe a Paul Samuelson, que la introdujo en la tercera edición de su manual (1955), y entre sus primeros arquitectos destacan tres[3]:

- *John Hicks*: formaliza las ideas de Keynes en el modelo IS-LM (*Mr. Keynes and the Classics*, 1937).

- *Alvin Hansen*: lo difunde y populariza, de ahí el nombre de modelo de Hicks-Hansen.

- *Franco Modigliani*: en 1944 lo dota de un mercado de trabajo y muestra que, salvo en el caso de la trampa de liquidez, el desempleo involuntario descansa, en última instancia, en la rigidez de los salarios nominales.

La idea que da nombre a la escuela es esa síntesis de dos visiones según el horizonte temporal. A corto plazo, la economía se comporta de forma keynesiana (rigideces, fallos de mercado y posibilidad de desempleo), y a largo plazo, de forma neoclásica (flexibilidad, vaciado de los mercados y pleno empleo). Sus fundamentos pueden ordenarse en cinco dimensiones, las mismas que cuestionará el monetarismo (apartado 2).

Metodología: la SNC formaliza matemáticamente las ideas de Keynes, cuya obra era esencialmente verbal[4], traduciéndolas a modelos manejables. Su núcleo es el modelo renta-gasto con mercado de trabajo (o IS-LM con mercado de trabajo), sobre el que se articulan después el modelo OA-DA y la curva de Phillips. La SNC combina, además, inducción y deducción, ya que impulsa la estimación econométrica de las grandes funciones de comportamiento (consumo, inversión, demanda de dinero) y los primeros grandes modelos macroeconométricos (p. ej. Klein), y hereda el método optimizador neoclásico para microfundamentarlas.

Supuestos:

- *Entorno*: el rasgo distintivo de la SNC es su carácter dual:

  - *A corto plazo* admite fallos de mercado, por lo que la economía no tiende automáticamente al pleno empleo. Entre ellos figuran la información imperfecta y la incertidumbre, que otorgan un papel central a las expectativas, aunque la SNC no ofrece un modelo explícito de su formación. El principal es la rigidez de los salarios nominales a la baja, que impide el vaciado del mercado de trabajo y hace posible el desempleo involuntario. Por ello, en este horizonte se rompe la dicotomía clásica y el dinero no es neutral.

  - *A largo plazo*, en cambio, la SNC mantiene el marco neoclásico de precios y salarios flexibles, vaciado de los mercados, pleno empleo y neutralidad del dinero. El mecanismo que reconcilia ambos planos es el efecto saldos reales (efecto Pigou-Patinkin), que devuelve la economía al pleno empleo una vez que precios y salarios se ajustan.

[3] La desarrollaron después, en los años cincuenta y sesenta, el propio Samuelson y Robert Solow, en el MIT, James Tobin, en Yale, y Lawrence Klein, con los primeros grandes modelos macroeconométricos (Klein y Goldberger, 1955).

[4] El diagrama renta-gasto (la cruz o aspa keynesiana) se debe a Samuelson (1948a).

▪ *Agentes*: como en el marco neoclásico, son optimizadores que persiguen su interés personal, pero a corto plazo no ajustan plenamente sus decisiones a las variables reales. En particular, los trabajadores padecen *ilusión monetaria*, un sesgo cognitivo por el que atienden al salario nominal y no al real[5]. Ello sustenta la rigidez salarial a la baja, pues se resisten a recortes nominales, pero aceptan caídas del salario real vía inflación.

Ajuste: es igualmente dual:

▪ *A corto plazo, ajuste por cantidades (mecanismo marshalliano)*: ante una caída de la demanda, los mercados responden con la producción y el empleo, y no con los precios. La variable de ajuste es la inversión no planeada, es decir, la acumulación o desacumulación de existencias. En coherencia con ello, la SNC niega (e invierte) la ley de Say[6]. Así, la demanda efectiva determina la oferta agregada, de modo que la economía puede estabilizarse en un equilibrio por debajo del pleno empleo (es decir, un equilibrio de subempleo).

▪ *A largo plazo, ajuste por precios*: la flexibilidad de precios y salarios restablece el pleno empleo, donde vuelven a regir los resultados neoclásicos (y se recupera la ley de Say).

Política económica: dado que a corto plazo los mercados no garantizan el pleno empleo, la SNC defiende un intervencionismo activo de gestión de la demanda (políticas fiscal y monetaria) para estabilizar el ciclo y cerrar la brecha de producción (*output gap*). Esta confianza en la sintonía fina de la demanda, y en la curva de Phillips como disyuntiva explotable, será el blanco principal de la crítica monetarista.

1.2. Análisis del lado de la demanda

Sobre estas bases, el modelo de la SNC parte de unos supuestos sencillos: corto plazo, economía cerrada y tres mercados (de bienes, de dinero y de bonos). El corto plazo implica precios fijos, de modo que la oferta agregada es perfectamente elástica al nivel de precios dado y se adapta pasivamente a la demanda, y el tipo de interés real coincide con el nominal. Su estudio se ordena en dos bloques. En primer lugar, el lado de la demanda determina el *output* (aspa keynesiana) y el tipo de interés (mercado de dinero). En segundo lugar, el lado de la oferta introduce el mercado de trabajo y la relación inflación-desempleo y relaja el supuesto de precios fijos.

1.2.1. Modelo IS-LM

Planteamiento: para la síntesis neoclásica, el comportamiento de la demanda se resume en el modelo IS-LM, un modelo de equilibrio general que condensa en dos ecuaciones el comportamiento de una economía cerrada (figura A5.3).

Curva IS: la primera de ellas es la *ecuación IS* (*Investment-Savings*), que viene dada por la siguiente fórmula:

$$\text{IS}: Y = C_0 + cY^d + I_0 - vr + G_0, \qquad c \in (0,1)$$

donde:

[5] Fisher (1928) dedicó a la ilusión monetaria su libro *The Money Illusion*, y es habitual asociarla al análisis keynesiano del mercado de trabajo, aunque el propio Keynes apela a la preocupación por los salarios relativos (tema A28, apartado 1.1.2).

[6] La SNC invierte la relación causal entre ahorro e inversión, es decir, la inversión determina el ahorro, y no al revés. Greenlaw y Shapiro (2017) llaman *ley de Keynes* al principio de que la demanda crea su propia oferta, inverso de la ley de Say.

- Y es la producción (o *output*) agregada de la economía.

- $C_0 + cY^d$ es la demanda de consumo agregada de la economía, que tiene un componente autónomo, C_0, y un componente inducido que depende de la renta, cY^d. En el componente inducido, $c \in (0, 1)$ es la propensión marginal a consumir de los hogares y $Y^d = Y - T_0$, su renta disponible después de impuestos de suma fija.

- $I_0 - vr$ es la demanda de inversión agregada de la economía, que tiene un componente autónomo, I_0, que recoge los *animal spirits* de los inversores, y un componente inducido, $-vr$, que recoge el efecto del tipo de interés real sobre la demanda de inversión.

- G_0 es un componente exógeno que recoge el gasto público (no productivo) realizado por el sector público.

La ecuación IS, que incorpora las funciones de comportamiento del gasto (consumo e inversión), es la condición de equilibrio en el mercado de bienes y servicios. La producción agregada (término de la izquierda) es igual al gasto agregado (término de la derecha).

Gráficamente, la derivación de la curva IS se representa en el plano renta-tipo de interés, con el aspa keynesiana en el panel superior (figura A5.1):

1. *Aspa keynesiana*: en el plano renta-demanda agregada, la demanda agregada para el tipo de interés i_0, $\text{DA}(i_0)$, corta a la recta de 45° en el punto 0, con renta Y_0. Con un tipo más bajo, i_1, la inversión es mayor y la demanda agregada, $\text{DA}(i_1)$, corta a la recta de 45° en el punto 1, con renta Y_1.

2. *Curva IS*: lleva los pares (Y_0, i_0) y (Y_1, i_1) al panel inferior. Es decreciente porque un tipo de interés más bajo eleva la inversión y, por el multiplicador, la renta de equilibrio.

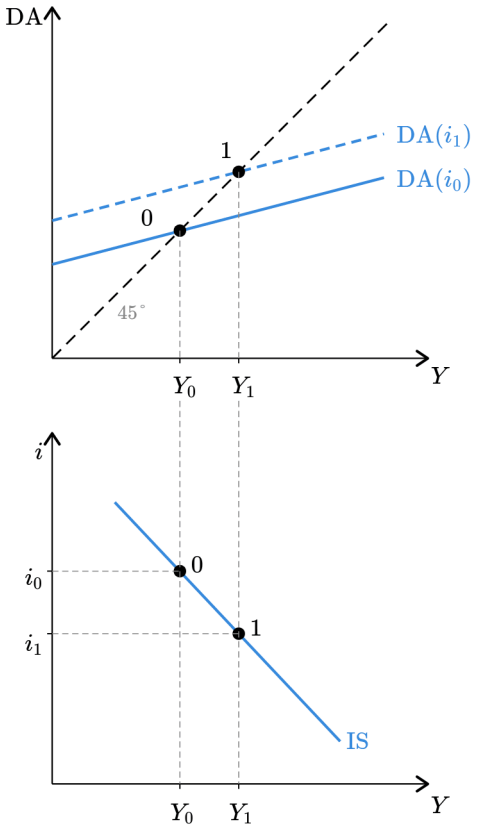

Figura A5.1: Derivación de la curva IS. La curva reúne las combinaciones de tipo de interés y renta que equilibran el mercado de bienes.

Curva LM: a la IS la acompaña una segunda ecuación de comportamiento conocida como *ecuación LM* (*Liquidity-preference-Money-supply*):

$$\text{LM}: \frac{M^s}{P} = L(Y^+, i^-) \rightarrow i = i(Y^+, \frac{M^s}{P}^-)$$

donde:

- M^s/P es la oferta monetaria real de la economía, la cual depende de las decisiones del banco central y la evolución de los precios de la economía (en el modelo IS-LM considerados constantes por ser un modelo de corto plazo).
- Y es la renta de los hogares.
- i es el tipo de interés nominal, que coincide con el real, $r = i - \pi^e$, cuando $\pi^e = 0$.

La ecuación LM, que incorpora la función de comportamiento de la demanda de dinero, es la condición de equilibrio en el mercado de dinero y bonos (mercado monetario). La oferta monetaria real, es decir, deflactada por el índice de precios (término de la izquierda), es igual a la demanda de dinero real (término de la derecha), que depende de la renta real de los hogares y del tipo de interés.

Respecto a la LM, es habitual presentar la ecuación en su forma lineal:

$$\text{LM}: \frac{M^s}{P} = kY - hi \rightarrow i = \frac{1}{h}(kY - \frac{M^s}{P})$$

Sin embargo, esta modelización no permitiría graficarla recogiendo el carácter asintótico horizontal de la curva en el plano (Y, i), que es la forma habitual en la que aparece en la mayoría de gráficos.

Gráficamente, la derivación de la curva LM se representa en el plano renta-tipo de interés, con el mercado de dinero en el panel izquierdo (figura A5.2):

1. *Mercado de dinero*: en el plano saldos reales-tipo de interés, la oferta monetaria real, M^s/P, es vertical y la demanda de dinero, decreciente en el tipo de interés. Con renta Y_0, la demanda $\frac{M^d}{P}(Y_0)$ fija el tipo i_0 (punto 0). Con una renta mayor, Y_1, la demanda se sitúa más a la derecha, $\frac{M^d}{P}(Y_1)$, y el tipo sube a i_1 (punto 1).

2. *Curva LM*: lleva los pares (Y_0, i_0) y (Y_1, i_1) al panel derecho. Es creciente porque una renta mayor eleva la demanda de dinero por motivo transacción y el tipo de interés sube para equilibrar el mercado.

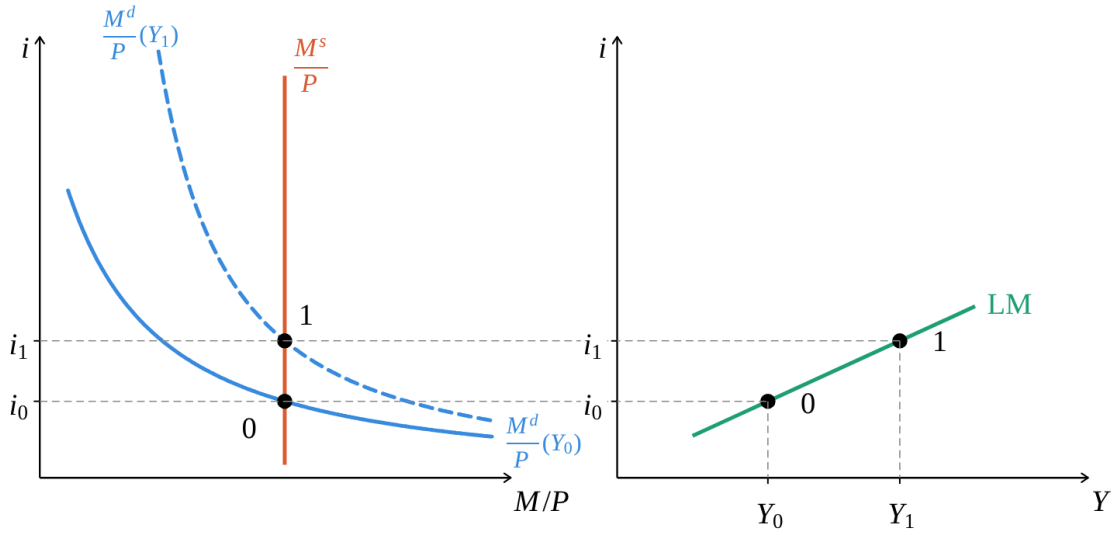

Figura A5.2: Derivación de la curva LM. La curva reúne las combinaciones de tipo de interés y renta que equilibran el mercado de dinero para una oferta monetaria real dada.

Equilibrio: de la intersección de ambas curvas se puede obtener la producción de equilibrio de la economía y el tipo de interés de equilibrio.

Desarrollo gráfico: el equilibrio del modelo IS-LM se representa en el plano renta-tipo de interés (figura A5.3):

1. *Curva IS*: recoge los pares de renta y tipo de interés que equilibran el mercado de bienes. Es decreciente porque un tipo de interés más bajo eleva la inversión y, por el multiplicador, la renta (figura A5.1).

2. *Curva LM*: recoge los pares que equilibran el mercado de dinero. Es creciente porque una renta mayor eleva la demanda de dinero y el tipo de interés que vacía el mercado (figura A5.2).

3. *Equilibrio*: en el punto 0, con renta Y^* y tipo de interés i^*, los dos mercados están a la vez en equilibrio.

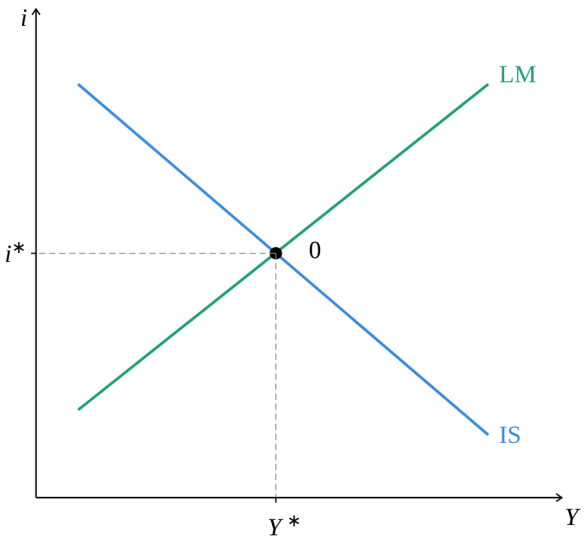

Figura A5.3: Equilibrio simultáneo en los mercados de bienes y de dinero en el modelo IS-LM.

Estática comparativa: con la IS y la LM en su forma lineal, y con $i = r$ (pues $\pi^e = 0$), la renta de equilibrio es

$$Y^* = \frac{h\,A_0 + v\,M^s/P}{h(1-c) + v\,k}, \qquad A_0 \equiv C_0 - cT_0 + I_0 + G_0$$

De esta expresión se deducen cuatro resultados:

- *Política monetaria*: un aumento de M^s/P desplaza la LM a la derecha, eleva la renta y reduce el tipo de interés, con un multiplicador $\partial Y/\partial(M^s/P) = v/[h(1-c) + vk]$.

- *Política fiscal*: un aumento de G_0 desplaza la IS a la derecha y eleva la renta y el tipo de interés. El multiplicador, $\partial Y/\partial G_0 = h/[h(1-c) + vk]$, es menor que el del aspa keynesiana, $1/(1-c)$, porque la subida del tipo de interés expulsa inversión privada, lo que se conoce como *efecto expulsión* (*crowding out*).

- *Pendientes*: en el plano (Y, i), la IS, de pendiente $-(1-c)/v$, es más vertical cuanto menor es v, y la LM, de pendiente k/h, es más plana cuanto mayor es h. La política fiscal gana eficacia cuanto más vertical es la IS y más plana la LM, y la monetaria en el caso contrario. Para la SNC, con una inversión poco sensible al tipo de interés (v bajo) y una demanda de dinero muy sensible a él (h alto), la política fiscal es la más eficaz[7] (anexo).

[7] Esa primacía de la política fiscal corresponde sobre todo a la primera síntesis, la de Hansen, y, en el Reino Unido, al Informe Radcliffe (1959), muy escéptico sobre la eficacia de la política monetaria. En los años sesenta, Tobin,

▪ *Casos extremos*: con una inversión insensible al tipo de interés ($v = 0$, IS vertical) o en la trampa de liquidez ($h \to \infty$, LM horizontal), el multiplicador fiscal alcanza el del aspa keynesiana, $1/(1-c)$, y la política monetaria no afecta a la renta.

1.2.2. Demanda de consumo

> La función de consumo keynesiana se expone en el tema A04, apartado 2.3.2, y se desarrolla, con los multiplicadores fiscales, la renta relativa y el ciclo vital, en el tema A33, apartados 1.2 y 2.2. Aquí se recoge lo que la síntesis tomó de estas teorías.

Teoría de la renta absoluta (Keynes, 1936): las premisas de Keynes son el consumo como fin último de la actividad económica, su dependencia de la renta corriente y una propensión marginal a consumir estable y menor que la unidad. Sobre ellas, los autores de la SNC construyen una función agregada de consumo lineal y estable en la renta real absoluta disponible del periodo:

$$C = C_0 + c\,(Y - T_0), \qquad C_0 > 0,\ c \in (0,1)$$

donde C_0 es el consumo autónomo, c la propensión marginal a consumir y T_0 un impuesto de suma fija. A diferencia de la visión clásica, el consumo depende sobre todo de la renta disponible del periodo y no del tipo de interés, y el ahorro es un residuo de la renta no consumida.

En una economía cerrada con impuestos de suma fija y tipo de interés dado (aspa keynesiana), los multiplicadores del gasto público y de los impuestos son $\partial Y/\partial G_0 = 1/(1-c) > 1$ y $\partial Y/\partial T_0 = -c/(1-c) < 0$. De ellos se siguen sus implicaciones de política económica:

▪ *Política monetaria*: es efectiva sobre el tipo de interés, pero, debido a los *animal spirits* y a la baja sensibilidad de la inversión al tipo de interés (es decir, el *eslabón débil*), apenas afecta a la renta y, por ello, al consumo.

▪ *Política fiscal*: será muy efectiva sobre el consumo, pues afecta a la renta del periodo y, a través del multiplicador, al consumo. Un aumento del gasto público es siempre más efectivo que una rebaja impositiva equivalente, porque la rebaja eleva la renta disponible, que los hogares reparten entre consumo y ahorro ($c < 1$). Por la misma razón, las políticas redistributivas estimulan el consumo, ya que los hogares de menor renta tienen una mayor propensión marginal a consumir.

Teoría de la renta relativa (Duesenberry, 1949): la utilidad de los hogares depende también del consumo de su entorno social (renta relativa, efecto demostración) y de sus niveles de consumo pasados (hábitos, efecto trinquete). Como el efecto trinquete hace el consumo rígido a la baja, en una recesión el consumo cae menos que la renta y la propensión media al consumo sube, lo que amortigua la caída de la demanda.

Hipótesis del ciclo vital (Modigliani y Brumberg, 1954; Ando y Modigliani, 1963): la hipótesis del ciclo vital (HCV), hoy ampliamente aceptada, gira en torno al papel que cumple el ahorro en el plan óptimo de consumo. Frente a una renta con un perfil previsible en forma de joroba, el consumo se mantiene estable, o crece ligeramente, durante toda la vida. Los hogares desplazan recursos de los años de ingresos altos a los de ingresos bajos (p. ej. la jubilación) y así suavizan su consumo.

Con la riqueza constante, la propensión media al consumo (PMeC) depende del horizonte:

Modigliani y Samuelson ya concedían a la política monetaria una eficacia considerable, de modo que el debate con los monetaristas pasó a ser una cuestión de grado, la del valor de los parámetros.

- *Corto plazo*: baja en las fases alcistas y sube en las bajistas, pues el aumento de renta se reparte entre varios periodos.

- *Largo plazo*: es estable, en línea con la evidencia de Kuznets.

Por la misma razón, una variación transitoria de la renta disponible se reparte entre los años de vida restantes y apenas altera el consumo corriente, de modo que la política fiscal transitoria es menos eficaz que en la función keynesiana (tema A33, apartado 2.2).

1.2.3. Demanda de inversión

> La eficiencia marginal del capital se desarrolla en el tema A34, apartado 1.2, y la teoría de la q de Tobin, en el apartado 2.2.2 del mismo tema. Aquí se recoge su papel en la síntesis.

Teoría de la eficiencia marginal del capital (Keynes, 1936): Keynes define la eficiencia marginal del capital (EMK) como la tasa de descuento que iguala el valor actual de los rendimientos esperados de un activo de capital con su precio de oferta (coste de reposición). Se invierte mientras la EMK supere al tipo de interés, de modo que la inversión es una función decreciente de este, $I = I_0 - v\,r$. Ahora bien, como la EMK depende de las expectativas sobre rendimientos futuros en un entorno de incertidumbre (los *animal spirits*), la inversión es muy volátil y poco sensible al tipo de interés (el eslabón débil). Por eso, para la SNC, la curva IS es muy inelástica al tipo de interés (casi vertical en el plano (Y, r)).

Teoría de la q (Tobin, 1969): ofrece una microfundamentación distinta de la demanda de inversión. La inversión óptima es función de la ratio q, entendida como el precio sombra relativo del capital (q marginal), que en la práctica se aproxima por la q media (el valor de mercado del capital instalado respecto de su coste de reposición). Se invierte cuando $q > 1$, de modo que la inversión (real) es una función creciente de q:

$$I_t = f(q_t), \qquad \partial I/\partial q > 0$$

donde I_t es la inversión y q_t la ratio q del periodo t. La teoría formaliza así la importancia de la estabilidad de los mercados financieros, ya apuntada por Keynes, al ligar la inversión real a las valoraciones de los activos.

1.2.4. Demanda de dinero

La teoría de la preferencia por la liquidez de Keynes distingue tres motivos para demandar dinero, transacción, precaución y especulación (tema A04, apartado 2.3.3), y la SNC desarrolló un modelo para microfundamentar cada uno de ellos.

> Los tres modelos (supuestos, programa de optimización, demanda de dinero y estática comparativa) se desarrollan en el tema A35, apartados 1.3.1 a 1.3.3. Aquí se recogen sus resultados.

Motivo transacción: modelo de inventarios (Baumol, 1952; Tobin, 1956): el dinero se mantiene como un inventario para transacciones del que se dispone de forma gradual. La gestión de ese inventario plantea una disyuntiva entre el coste de transacción (de convertir otros activos en dinero) y el coste de oportunidad de mantenerlo. La regla de la raíz cuadrada resultante hace

que la demanda real de dinero dependa positivamente del coste de transacción b, negativamente del tipo de interés nominal i y positivamente de la renta elevada a un medio:

$$M^d/P = k(b^+, i^-) \cdot Y^{1/2}$$

donde M^d/P es la demanda real de dinero e Y la renta real. La demanda de dinero es, por tanto, menos que proporcional a la renta (su elasticidad-renta es ½, lo que refleja economías de escala en la gestión de tesorería). Por ello, un aumento de la cantidad de dinero desplaza la LM más que si la demanda fuera proporcional a la renta, pues, a tipo de interés dado, la renta que equilibra el mercado de dinero responde de forma amplificada (en términos del modelo, $Y \propto (M^s/P)^2$), aunque su efecto final sobre la renta depende de la pendiente de la IS (apartado 1.2.1).

Motivo precaución: modelos de ruptura (Whalen, 1966): el dinero se mantiene como inventario frente a la incertidumbre sobre el flujo de entradas y salidas de caja y los gastos fortuitos (el riesgo de quedarse sin liquidez, es decir, la *ruptura*). El saldo precautorio óptimo equilibra el coste de oportunidad de mantenerlo con el coste esperado de las rupturas y sigue una regla de la raíz cúbica:

$$M^d = \left(\frac{2\sigma^2 b}{i}\right)^{1/3}$$

donde σ^2 es la varianza de los pagos netos, b el coste de cada ruptura e i el tipo de interés nominal. Como la varianza de los pagos netos crece con el volumen de transacciones y este, con la renta, el efecto de la renta es positivo (si la varianza es proporcional a la renta, la elasticidad-renta es 1/3), de modo que $M^d = f(b^+, i^-, Y^+)$.

Motivo especulación: modelo de cartera (Tobin, 1958): aplica la selección de cartera a la demanda especulativa y supone que los agentes diversifican entre dinero y bonos. Por ello, la demanda de dinero depende positivamente de la riqueza W y del riesgo de los bonos σ_B, y de forma ambigua de su rentabilidad r_B:

$$M^d = f(W^+, \sigma_B^+, r_B^\pm)$$

donde M^d es la demanda de dinero. En consecuencia, la política monetaria actúa alterando la rentabilidad relativa de los bonos y, con ella, la composición de las carteras. Su efecto sobre la renta es menor cuanto más sensible sea la demanda de dinero al tipo de interés.

1.2.5. Conclusiones

El análisis de la demanda muestra cómo la SNC fue más allá de las funciones de comportamiento de Keynes, microfundamentándolas sobre el comportamiento optimizador de los agentes.

Aspectos positivos:

- *Microfundamentación*: dota a los componentes de la demanda de una base optimizadora de la que carecía Keynes. Por ejemplo, el consumo se orienta a los recursos vitales (HCV), la inversión se liga a las valoraciones de los activos (teoría de la q) y la demanda de dinero gana rigor (inventarios y carteras).

- *Capacidad explicativa*: resuelve regularidades empíricas como la paradoja de Kuznets (la divergencia entre la PMeC de corto y de largo plazo), mediante la renta relativa y la HCV, y conecta la inversión con los mercados financieros.

- *Marco para la política*: ofrece un cuadro coherente de gestión de la demanda (multiplicadores y componentes del IS-LM), acorde con la defensa de la estabilización a corto plazo propia de la síntesis.

Limitaciones:

- *Ausencia de una teoría de las expectativas*: las funciones de demanda no explican cómo se forman las expectativas, laguna que el monetarismo (con la hipótesis de expectativas adaptativas, HEA) y la nueva macroeconomía clásica (NMC, con la hipótesis de expectativas racionales, HER) tratarán de cubrir.

- *Fragilidad del "eslabón débil"*: hacer descansar la transmisión en una inversión poco sensible al tipo de interés (dominada por los *animal spirits*) y en un consumo guiado por la renta corriente debilita la política monetaria. Los monetaristas rebatirán ese supuesto (inversión sensible al tipo de interés y demanda de dinero estable).

- *Segmentación de la demanda de dinero*: la división en tres motivos resulta en parte *ad hoc*, y Friedman tratará el dinero como un activo más, con una demanda unificada.

- *Separación de decisiones*: el consumo-ahorro y la inversión se analizan por separado (apartado 2.1).

1.3. Análisis del lado de la oferta

En términos de política económica, el paradigma keynesiano buscaba dar con un nivel adecuado de demanda agregada, pues su falta provocaba desempleo involuntario y su exceso, inflación. Sin embargo, hasta los años cincuenta los estudios para fijar ese nivel estaban dispersos, y la determinación de la inflación y su relación con el mercado de trabajo era el punto débil del keynesianismo.

1.3.1. Relación entre inflación y desempleo: modelo de la curva de Phillips (Phillips, 1958; Lipsey, 1960)

Origen: durante la década de 1950, el economista neozelandés Alban Phillips estudiaba en la London School of Economics (LSE) el marco keynesiano y cómo, según Keynes, en las recesiones coincidían un desempleo involuntario alto y unas presiones inflacionistas bajas. Con datos del Reino Unido entre 1861 y 1957, Phillips comprobó que los salarios nominales crecían más despacio (menor inflación salarial) cuanto mayor era la tasa de desempleo, y que esa relación no era lineal[8]. El buen ajuste sugería, además, que el vínculo era estable y duradero[9]. A raíz de su obra surgieron dos líneas de investigación, una empírica, que buscaba confirmar la estabilidad de la relación, y otra teórica, que buscaba explicarla a partir del funcionamiento del mercado de trabajo.

Idea: inflación y paro varían en sentido opuesto, siguiendo una curva estable y no lineal, de modo que la política económica puede elegir entre más inflación y menos paro, o a la inversa.

[8] Phillips añade que la relación es muy convexa (los salarios nominales apenas bajan aunque el paro sea alto) y que también influye la velocidad de variación del paro, de modo que, si el desempleo cae deprisa, la inflación salarial se acelera.

[9] La relación tenía precedentes. Irving Fisher (1926) ya había hallado, con datos de Estados Unidos, una relación inversa entre las variaciones de los precios y el desempleo, aunque para él la causalidad iba de los precios al empleo. Su artículo se reeditó en 1973 con el título *I Discovered the Phillips Curve*.

El fundamento teórico de Lipsey (sus tres supuestos y la derivación de la curva) se desarrolla en el tema A28, apartado 1.2.1. Aquí se recoge su papel en la síntesis neoclásica.

Razonamiento: Richard Lipsey (1960) dio a la relación su fundamento teórico[10]. Cuanto más alto es el paro, más pequeña es, de forma no lineal, la demanda excedente de trabajo y más despacio suben los salarios nominales. A su vez, con productividad constante, los precios suben al mismo ritmo que los salarios (Samuelson y Solow, 1960). El resultado es la *curva de Phillips*, que liga de forma inversa y convexa la inflación con el desempleo:

$$\pi_t = f(u_t), \qquad f' < 0,\ f'' > 0$$

donde π_t es la tasa de inflación y u_t la tasa de desempleo. Dicho de otro modo, existe una disyuntiva (*trade-off*) entre inflación y desempleo.

Desarrollo gráfico: en el plano desempleo-inflación (figura A5.4), la curva original es decreciente y convexa, y puede adoptar distintas formas según su pendiente, entre dos casos polares:

1. *Curva vertical*, $\text{CP}_{\text{polar clásica}}$: a un mismo nivel de desempleo (el natural, $\bar{u}$) puede corresponder cualquier inflación, que es así un fenómeno de demanda (monetario), caso que suele asociarse a la visión clásica.

2. *Curva horizontal*, $\text{CP}_{\text{polar keynesiana}}$: a cualquier nivel de desempleo corresponde una única inflación, $\bar{\pi}$, determinada por los costes, que es así un fenómeno de oferta, caso que suele asociarse a la visión keynesiana más estricta.

Ninguna de las dos escuelas sostuvo, sin embargo, una forma exactamente polar, y ambos casos son simplificaciones que sirven para ordenar el debate.

[10] Lipsey obtenía primero curvas de Phillips por industria y las agregaba después, de modo que la inflación no solo depende del nivel de desempleo de la economía, sino también de su distribución entre industrias.

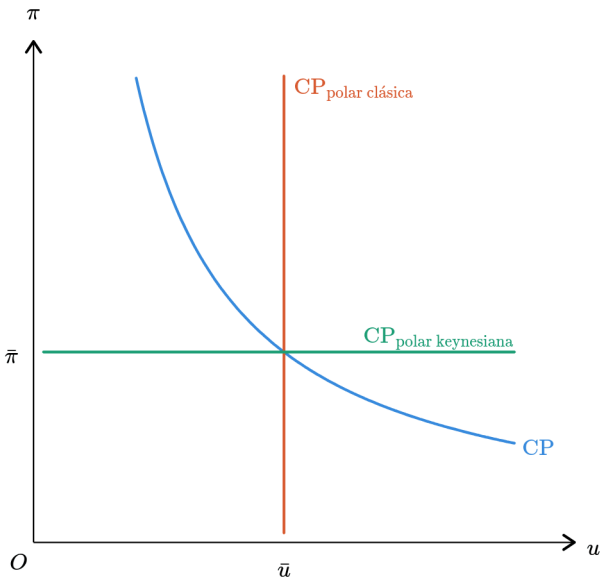

Figura A5.4: Formas de la curva de Phillips. La curva original es decreciente y convexa. Sus dos casos polares son la vertical en el desempleo natural, $\bar{u}$, que suele asociarse a la visión clásica, y la horizontal en la inflación de costes, $\bar{\pi}$, que suele asociarse a la keynesiana.

Extensiones: Samuelson y Solow (1960) dedujeron, además, que, si la curva de Phillips era empíricamente estable, los responsables de la política económica (*policy-makers*) podían elegir una tasa de desempleo concreta a cambio de una determinada inflación. Ambos autores concibieron así la curva como un *menú* de política económica que permitía un "ajuste fino" del mercado laboral, es decir, mantener el desempleo bajo a cambio de una inflación continuada mediante políticas monetarias expansivas sostenidas. Advirtieron, con todo, de que ese menú podía no mantener su forma a más largo plazo[11].

Conclusiones:

▪ *Aspectos positivos*: la curva condensaba en una única relación el nexo entre el mercado de trabajo y la inflación, sin necesidad de microfundamentar las rigideces laborales y su traslado a los precios. Por ello fue, durante los años sesenta, el instrumento central de la síntesis neoclásica en el análisis del mercado de trabajo y, por extensión, de la oferta agregada.

▪ *Limitaciones*: la relación no resultó estable, porque con la estanflación de los años setenta subieron a la vez la inflación y el desempleo. Además, el razonamiento de Lipsey no partía de agentes que optimizan, de modo que la curva carecía de microfundamentos en sentido estricto y exigía ilusión monetaria permanente (apartado 2.3.1).

[11] A corto plazo, cualquier medida de política de demanda produce movimientos a lo largo de la curva en el plano desempleo-inflación, y no desplazamientos de esta, ya que las variables que se alteran (p. ej. masa monetaria, gasto público, renta…) no aparecen explícitamente en ella. Lo que sí producen es un desplazamiento de la curva de DA y, con él, un movimiento a lo largo de la curva de OA en el plano renta-precios.

1.3.2. Relación entre producción y desempleo: la ley de Okun (Okun, 1962)

Origen: casi en paralelo al desarrollo de la curva de Phillips, el economista estadounidense Arthur Okun detectó empíricamente que la tasa de paro y la tasa de crecimiento de la producción estaban relacionadas.

Idea: sus estimaciones con las series de Estados Unidos entre 1947 y 1960 apuntaban a que un aumento de un punto en la tasa de paro se asociaba a una pérdida de producción de en torno al 3 %. Esta respuesta es mucho mayor que la que predice una función de producción con productividad marginal del trabajo decreciente, porque al empleo se suman los cambios en las horas trabajadas, en la participación laboral y en la productividad (atesoramiento de mano de obra).

Razonamiento: la relación, bautizada después como *ley de Okun*, es, pese a su simplicidad, una herramienta útil de política económica. En una de sus versiones originales, en primeras diferencias, plantea una relación de proporcionalidad (inversa) entre la desviación del crecimiento del *output* respecto de su tasa tendencial y la variación de la tasa de desempleo:

$$u_t - u_{t-1} = -\frac{1}{\phi}\,(g_{Y,t} - \bar{g}_Y), \qquad \phi > 0$$

donde ϕ es el coeficiente de Okun (en torno a 3 en sus estimaciones) y $\bar{g}_Y$ es la tasa de crecimiento tendencial del *output*. La producción debe crecer a ese ritmo solo para mantener el desempleo constante, y un crecimiento superior lo reduce.

La ley admite también una formulación en brechas, que relaciona el desempleo cíclico con la brecha de producción:

$$y_t - \bar{y}_t = -\phi\,(u_t - \bar{u})$$

donde $y_t \equiv \ln Y_t$ y $\bar{y}_t$ es la producción potencial. Si la tasa natural $\bar{u}$ es constante, la diferencia de un periodo a otro devuelve la versión anterior, de modo que ambas recogen la misma regularidad. La formulación en brechas es, además, la que traslada la curva de Phillips al plano producción-inflación. En su versión lineal, $\pi_t = -\kappa_u(u_t - \bar{u})$, basta sustituir el desempleo cíclico para obtener $\pi_t = \kappa\,(y_t - \bar{y}_t)$, con $\kappa = \kappa_u/\phi > 0$, una oferta agregada con pendiente positiva que cierra el modelo IS-LM por el lado de la oferta (tema A28, apartado 1.2.2).

Conclusiones:

- *Aspectos positivos*: la ley tuvo gran impacto en EE. UU., donde las estimaciones oficiales de la producción potencial y del superávit presupuestario asociado se convirtieron en indicadores esenciales para graduar las políticas de estabilización. Pese a su naturaleza "a-teórica" y al debate sobre su estabilidad, se considera una de las regularidades empíricas más sólidas.

- *Limitaciones*: Okun era consciente de que sus ecuaciones eran representaciones muy simples de interacciones complejas, con riesgo de sobreestimación. En cuanto al coeficiente de Okun, se observa una notable disparidad entre países y una ruptura estructural en torno a 1973-1974, si bien la ley parecía resistir mejor que la curva de Phillips tradicional.

1.3.3. Conclusiones

Aspectos positivos: las dos relaciones cubrieron el punto débil del keynesianismo, la determinación de la inflación y su relación con el mercado de trabajo, y completaron así el modelo IS-LM con un análisis de la oferta.

Limitaciones: a finales de los años sesenta ese análisis entra en crisis por razones teóricas y empíricas. Durante Bretton Woods (1944-1971), y en especial desde los años sesenta, el paro era bajo en las economías desarrolladas. A partir de 1970 esa tendencia se revierte y aparece la *estanflación*, es decir, la coincidencia de alto desempleo e inflación, de modo que la curva de Phillips deja de ser estable. La curva original no podía recoger esta inflación de oferta, en la que el aumento de la inflación se acompaña de más desempleo, y no de menos.

Según Michael Bruno y Jeffrey Sachs (1985), la estanflación se debió a la conjunción de cuatro elementos que dificultaban el ajuste en un contexto de inflación elevada:

- *Perturbaciones de oferta*: el encarecimiento de las materias primas y, sobre todo, del petróleo (1973-1974 y 1979).

- *Productividad*: la caída del crecimiento de la productividad total de los factores.

- *Rigidez del salario real*: los salarios reales no se ajustaron a la baja tras el encarecimiento de los insumos, lo que abrió una *brecha salarial*, sobre todo en Europa.

- *Respuesta de la política económica*: la contracción de la demanda con la que los Gobiernos y los bancos centrales combatieron la inflación.

1.4. Conclusiones

Aspectos positivos: la síntesis neoclásica dio a la macroeconomía un instrumental común que hizo posible tanto la enseñanza sistemática de la disciplina como la estimación econométrica de sus grandes funciones de comportamiento. Su carácter dual, con precios rígidos a corto plazo y flexibles a largo, permitió además reconciliar la intervención keynesiana con el equilibrio neoclásico sin renunciar a ninguno de los dos.

Limitaciones: esa misma dualidad era su punto débil, pues la rigidez de precios y salarios se postulaba sin microfundamentar. La ausencia de una teoría de la formación de expectativas dejó también el modelo expuesto a la crítica. Cuando la estanflación de los años setenta desmintió la estabilidad empírica de la curva de Phillips, la síntesis carecía de recursos internos para explicar la coexistencia de inflación y desempleo elevados. Ese doble fallo, empírico y teórico, propició la contrarrevolución monetarista y, más tarde, la nueva macroeconomía clásica.

2. Monetaristas

Estas limitaciones dieron paso a la crítica monetarista[12], en la que se aprecian dos etapas. La primera (de mediados de los cincuenta a finales de los sesenta) se centró en el lado de la demanda, y la segunda (desde finales de los sesenta), en el lado de la oferta. Su prolongación en la economía abierta es el enfoque monetario de la balanza de pagos (Johnson, 1958, 1972), que traslada la misma lógica al sector exterior. Con tipo de cambio fijo, el saldo de la balanza de pagos refleja el exceso de oferta o de demanda de dinero, el ajuste es automático y la política monetaria pierde autonomía (tema B12, apartado 2.1).

2.1. Principales rasgos

Los rasgos del monetarismo se ordenan en las mismas dimensiones que los de la síntesis neoclásica (apartado 1.1).

Metodología: los monetaristas se apoyan en las herramientas del marco IS-LM. No obstante, Friedman se separa de Keynes y de muchos keynesianos, que reprochaban a los neoclásicos lo poco realistas que eran sus supuestos. Para Friedman ese realismo es irrelevante, porque una teoría se juzga por sus predicciones. En esta *tesis de irrelevancia de los supuestos* (Friedman, 1953)[13] se apoya en gran parte el pensamiento monetarista.

Supuestos:

▪ *Entorno*: parten de muchos de los postulados (neo)clásicos originales (p. ej. racionalidad y flexibilidad de precios y salarios). Sin embargo, no defienden la dicotomía clásica, porque para ellos el dinero es neutral a largo plazo, pero no a corto. Esta idea se resume en el lema monetarista *money matters* ("el dinero importa"). Por tanto, la política monetaria altera las variables reales durante el corto plazo. En ese horizonte, los ciclos económicos se deben sobre todo a las oscilaciones de la oferta monetaria, motor de la demanda (la causalidad va del dinero a la demanda, no al revés).

▪ *Agentes*: una de las principales limitaciones de la SNC era la ausencia de una teoría sobre la formación de expectativas. Para suplirla, la escuela monetarista popularizó la *hipótesis de expectativas adaptativas* (HEA) anticipada por Irving Fisher (1911) y formalizada por Cagan (1956) y Nerlove (1958). La HEA fue fundamental en la crítica a la curva de Phillips y para estimar la renta permanente en la teoría del consumo de Friedman (hipótesis de la renta permanente, HRP). Formalmente, la HEA admite distintas especificaciones, si bien generalmente se utiliza la versión de retardos distribuidos con pesos geométricos, $X^e_{t+1} = \lambda \sum_{j=0}^{\infty} (1-\lambda)^j X_{t-j}$. En ella, X^e_{t+1} es el valor esperado de la variable para el periodo siguiente, X_{t-j} sus valores pasados observados y $\lambda \in (0, 1]$ el parámetro de ajuste, que pondera los valores pasados con pesos geométricamente decrecientes. Así, cada periodo la expectativa se revisa en una fracción λ del último error de previsión.

Ajuste: el pensamiento keynesiano había dado lugar a dos variables de ajuste, las existencias o inversión no planeada en el mercado de bienes y servicios, y el tipo de interés (real) en el mercado monetario. En consecuencia, existía una separación entre las decisiones de consumo y ahorro y las decisiones de inversión. Friedman rompe esta separación e introduce el mecanismo

[12] Además de Milton Friedman y Anna Schwartz, la escuela tuvo otra pareja de referencia en Karl Brunner, que le dio nombre, y Allan Meltzer, y en el Reino Unido destacaron David Laidler y Michael Parkin.

[13] Samuelson (1963) la llamó *F-twist* y objetó que unos supuestos irreales no se vuelven aceptables porque algunas de sus predicciones se cumplan.

de transmisión monetarista, que abre un canal directo de ajuste a través de la riqueza de los individuos, vía consumo y demanda de dinero.

Política económica: sus recomendaciones chocan, por lo general, con las de la SNC prácticamente punto por punto. La política fiscal es mucho menos eficaz que la monetaria, porque el consumo, que depende de la renta permanente, es estable y la demanda de dinero, al tener el dinero muchos sustitutivos imperfectos, es estable y poco sensible al tipo de interés. De ahí que el gasto público expulse gasto privado (efecto expulsión). Para estabilizar la economía, en cambio, ningún instrumento iguala a la política monetaria, aunque debe aplicarse con reglas y no a discreción. Por ello, Friedman recomienda una regla de crecimiento constante de un agregado monetario, la *regla del k %* (1960).

2.2. La crítica por el lado de la demanda

En este primer frente, Friedman y sus seguidores reconstruyen tres funciones de la síntesis (consumo, inversión y demanda de dinero) y deducen de ellas sus consecuencias para la política económica.

2.2.1. Demanda de consumo monetarista: hipótesis de la renta permanente (Friedman, 1957)

Origen: Friedman (1957) construye una teoría de la demanda de consumo influida por las obras de Arthur Pigou e Irving Fisher.

Idea: el consumo privado no depende de la renta del periodo como tal, sino de la *renta permanente*, es decir, del rendimiento de la riqueza total (humana y no humana).

> La hipótesis de la renta permanente (supuestos, programa del consumidor, funciones de consumo y estática comparativa) se desarrolla en el tema A33, apartado 2.3. Aquí se recogen los resultados en que se apoya la crítica monetarista.

Planteamiento: el modelo consta de dos bloques, la renta y el consumo, cada uno con un componente permanente y otro transitorio:

- *Renta*: la renta corriente se divide en dos componentes no correlacionados, $Y_t = Y_t^P + Y_t^T$. Lo máximo que el consumidor puede gastar en cada periodo sin reducir su riqueza real neta total, humana y no humana, es su *renta permanente*, Y_t^P, que se obtiene al multiplicar esa riqueza por r, un tipo real de interés a largo plazo:

$$Y_t^P = rW_t, \qquad W_t = \frac{B_t}{r} + M_t + K_t + A_t + W_H, \qquad W_H = \sum_{s=t}^{\infty} \frac{1}{(1+r)^{s-t}} y_s^e$$

donde B_t/r es el valor real de los bonos (perpetuos), M_t los saldos monetarios reales, K_t los activos físicos (p. ej. la vivienda), A_t las acciones y W_H la riqueza humana, valor descontado de las rentas laborales esperadas y_s^e. Como la renta permanente no es observable, los consumidores, previsores (*forward looking*), la estiman con expectativas adaptativas (apartado 2.1). La *renta transitoria*, Y_t^T, es la desviación aleatoria (positiva o negativa)[14] de la renta corriente respecto a la permanente (p. ej. lotería).

[14] En el agregado, las rentas positivas y negativas no se cancelan en su totalidad. Por ello, siempre alguna parte de la renta agregada es transitoria (Argandoña, Gámez y Mochón, 1996).

▪ *Consumo*: se descompone del mismo modo en dos partes sin correlación entre sí. El *consumo permanente* guarda una relación proporcional[15] con la renta permanente, $C_t^P = k_P Y_t^P$, con $k_P = f(r^-, (W_H/W)^-, z^\pm) \in (0,1)$. Reducen k_P tanto un tipo de interés más alto, que encarece el consumo presente, como un mayor peso de la riqueza humana, más arriesgada y peor aval ante un prestamista. Otros factores, z (gustos, demografía, incertidumbre), lo mueven en uno u otro sentido, y en cambio no le afecta cuánta renta permanente se tenga. El *consumo transitorio* es un componente aleatorio (p. ej. un gasto por enfermedad) que tampoco guarda correlación con la renta transitoria:

$$\rho(C_t^T, Y_t^T) = 0$$

La propensión marginal a consumir la renta transitoria es, por tanto, nula, de modo que un *shock* de renta transitoria positivo se ahorra íntegramente y uno negativo (p. ej. un despido por la COVID) se absorbe con desahorro.

Solución: la función de consumo agregada depende del horizonte:

▪ *Largo plazo*: los *shocks* transitorios de consumo se cancelan en el agregado y la renta transitoria ahorrada pasa a formar parte de la permanente[16], de modo que $C_t = k_P Y_t^P$ y la propensión media al consumo es constante, en línea con la evidencia de Kuznets.

▪ *Corto plazo*: como la renta permanente se estima como un retardo distribuido de la renta corriente, la función de consumo se aproxima con una keynesiana, $C_t = C_0 + c\,Y_t$, aunque los agentes consumen según su renta permanente y ajustan el ahorro.

Ello explica que la propensión media al consumo de los hogares de renta baja sea, a corto plazo, mayor que k_P (y que la de los hogares de renta alta) y que la PMeC sea anticíclica (tema A33, anexo). La renta permanente fundamenta, además, el mecanismo de transmisión monetarista.

Extensiones: el fenómeno del alisamiento o suavización del consumo derivado de la consideración de la riqueza en las decisiones de consumo hace que el comportamiento del consumo privado sea, en general, más estable que en Keynes. De ello se siguen sus implicaciones de política económica:

▪ *Política monetaria*: un aumento no anticipado transitorio de la oferta monetaria tiene un efecto expansivo sobre el consumo, pues provoca un efecto riqueza directo vía saldos monetarios e indirecto vía tipo de interés[17]. Sin embargo, si el aumento es permanente (es decir, se mantiene a largo plazo), termina siendo neutral sobre el consumo y provoca inflación, porque se cumple la teoría cuantitativa del dinero.

▪ *Política fiscal*: sus efectos dependen de la naturaleza del cambio:

- *No anticipado y transitorio*: afecta a la renta transitoria, pero no a la renta permanente. Por tanto, el ingreso adicional será plenamente ahorrado por los consumidores y no tendrá efecto sobre el consumo.

[15] Siempre y cuando la función de utilidad intertemporal sea homotética (Argandoña, Gámez y Mochón, 1996).
[16] Friedman estimó que la renta permanente solo acababa absorbiendo un cambio de la renta transitoria si este se mantenía durante tres años.
[17] Hay que tener en cuenta que el efecto riqueza indirecto sobre el consumo de una caída del tipo de interés en la HRP es siempre positivo, dado que $k_P = f(r^-, \cdot)$.

- *No anticipado y permanente*: el *shock* de renta sí se incorpora a la renta permanente y eleva el consumo. Que lo compense la anticipación de los impuestos futuros necesarios para financiarlo no es un resultado de Friedman, sino de Barro (1974), que lo formaliza como equivalencia ricardiana (tema A06, apartado 3.2.1).

Conclusiones:

- *Aspectos positivos*:

- *Compatibilidad*: encaja plenamente con la teoría del ciclo vital, y ambas llegan a predicciones similares.

- *Evidencia empírica*: explica, junto con la HCV, las regularidades observadas por Kuznets.

- *Limitaciones*:

- *Entorno*: supone mercados de capitales perfectos, en los que los hogares pueden endeudarse sin restricciones de liquidez contra su renta futura para suavizar el consumo, y que los hogares son capaces de estimar con fiabilidad su renta permanente.

- *Agentes*: resulta contradictorio que el consumidor intente suavizar el consumo durante todo su horizonte vital con una mirada hacia el futuro (*forward looking*) y que, en cambio, estime su Y^P mirando hacia atrás con expectativas adaptativas, que conducen a errores sistemáticos.

- *Contraste de hipótesis*: resulta difícil en la medida en que la renta permanente y el consumo permanente no son observables.

2.2.2. Demanda de inversión: diferentes sensibilidades

En cuanto a la demanda de inversión, en lo esencial Friedman acepta la teoría sobre la eficiencia marginal del capital de Keynes, en la medida en que esta, a su vez, se basaba en la "segunda aproximación" de Fisher. El debate giraba, por tanto, no tanto en torno a la forma funcional de la demanda de inversión como en torno a su sensibilidad al tipo de interés, es decir, al valor de v en la función de inversión de la síntesis, $I = I_0 - vr$.

Keynesianismo: para la síntesis neoclásica, ningún componente de la demanda agregada es tan volátil como la inversión, que se mueve al compás de los *animal spirits* y apenas reacciona al tipo de interés (el eslabón débil, apartado 1.2.3). Como, además, el consumo depende principalmente de la renta disponible y no del tipo de interés real, la curva IS es rígida al tipo de interés real, casi vertical en el plano renta-tipo de interés.

Monetarismo: los monetaristas conservan la función de inversión de la SNC, pero no su lectura. Los agentes reparten su riqueza entre muchos activos y se guían por su rentabilidad, de modo que la inversión oscila mucho menos, reacciona con fuerza a r y desaparece el eslabón débil. El consumo refuerza el resultado, pues depende de la riqueza (vía renta permanente), cuyo valor cambia con el tipo de interés. Por ambas vías, la IS va a ser elástica al tipo de interés (real).

2.2.3. Demanda de dinero monetarista: el dinero como bien (Friedman, 1956)

Origen: Milton Friedman (1956) reformula la teoría cuantitativa como una teoría de la demanda de dinero, que amplía la aportación de la SNC al tratar el dinero como un activo más.

Idea: los agentes económicos van a demandar una mayor variedad de activos financieros, no solo bonos, y el dinero es un bien más[18].

> La demanda de dinero de Friedman (supuestos, función, estática comparativa y valoración) se desarrolla en el tema A35, apartado 1.4.1. Aquí se recoge lo que sostiene la crítica monetarista.

Supuestos:

- *Entorno*: Keynes atribuía al dinero pocos sustitutos, aunque muy próximos. Friedman, en cambio, le reconoce muchos, todos imperfectos (p. ej. bonos y acciones, bienes de consumo duradero e incluso activos inmobiliarios[19]).

- *Agentes*: el dinero es un bien de capital más, factor de producción para las empresas y fuente de poder adquisitivo general para los consumidores, que rinde servicios y, en su caso, el interés de los depósitos. Frente a la separación de Keynes y la SNC, en Friedman la naturaleza del dinero como bien permite tratar de manera conjunta las decisiones de consumo-ahorro y las de inversión o cartera (Froyen, 2013, cap. 9). Friedman acepta que puedan existir distintos motivos para demandar dinero, pero no segmenta su demanda en componentes.

Razonamiento: al igual que Tobin, Friedman obtiene una función de demanda (real) de dinero mediante un enfoque de decisión de cartera[20]:

$$M^d/P = f((Y^P)^+, (W_H/W)^+, r_B^-, r_E^-, \pi^{e-})$$

$$M^d/P = k(r^-, \pi^{e-}) \cdot (Y^P)^+ \tag{1}$$

donde Y^P es la renta permanente, W_H/W es el peso de la riqueza humana (renta salarial capitalizada) en la riqueza total, r_B es la tasa de rentabilidad (nominal) de los bonos, r_E es la tasa de rentabilidad (nominal) de las acciones y π^e es la tasa de inflación esperada.

Agrupando las tasas de rentabilidad de los distintos activos y la riqueza se obtiene la ecuación (1). Esta expresión muestra que la teoría de la demanda de dinero es una reformulación de la ecuación de Cambridge en la que k depende de la tasa de rentabilidad de los activos financieros y reales. De ahí se siguen los dos resultados principales de la teoría monetarista:

- *Demanda de dinero estable*: como el dinero tiene un alto número de sustitutivos y su demanda depende de la renta permanente, la sensibilidad de esa demanda al tipo de interés es baja ($h \approx 0$).

[18] Keynes ya había puesto énfasis en el papel del dinero como activo.
[19] En estos bienes, el rendimiento está implícito y es en especie.
[20] La especificación puede incluir además la rentabilidad (nominal) del dinero, r_m (como los intereses de los depósitos), con $\partial M^d/\partial r_m > 0$, y un parámetro z de preferencias. A veces se expresa en rentabilidades (nominales) esperadas.

▪ *Mecanismo de transmisión*: la transmisión monetarista descansa en el reajuste de las carteras. Un exceso de saldos monetarios se gasta en una gama amplia de activos financieros y reales, cuyos precios suben y cuyos rendimientos bajan, hasta alcanzar el gasto corriente. La inflación esperada interviene como el rendimiento de los bienes frente al dinero, es decir, como un coste de mantener saldos, y liga así la demanda (real) de dinero con la de activos reales.

Conclusiones:

▪ *Aspectos positivos*: recupera la estabilidad de la función de demanda de dinero con un fundamento microeconómico, la elección de cartera, y hace de esa función el pilar de la teoría cuantitativa revisada y del mecanismo de transmisión monetarista.

▪ *Limitaciones*: no deriva la función de un programa de optimización explícito, sino que postula sus argumentos, y la estabilidad que le atribuye, tesis central del monetarismo, se quebró con la innovación financiera de los años ochenta.

2.2.4. Implicaciones de la microfundamentación monetarista de la demanda

Equilibrio (Friedman, 1956): a partir de su teoría de la demanda de dinero, Friedman lleva a cabo una reformulación de la teoría cuantitativa del dinero que algunos autores denominan *teoría cuantitativa revisada*.

La teoría parte de la ecuación de cambio, $MV = PY$, una identidad contable, y de su lectura causal neoclásica (de izquierda a derecha). Aunque Friedman cuestiona que la oferta monetaria sea exógena debido a la amplitud del agregado, acepta este supuesto como simplificación.

Combinando esta identidad con su demanda de dinero, deduce que la velocidad de circulación del dinero es muy estable, aunque no plenamente constante. La estabilidad viene de que, por la ecuación (1), $V = Y/(kY^P)$, la inversa de k cuando $Y = Y^P$, es una función estable de pocas variables (los rendimientos relativos de los activos y la ratio Y/Y^P), lo que la hace predecible. Friedman considera, además, la condición de equilibrio instantánea por la cual la oferta de dinero es igual a la demanda de dinero ($M^s = M^d$).

Finalmente, construye una versión dinámica de la teoría cuantitativa del dinero que rompe la dicotomía clásica en el corto plazo. Tomando logaritmos en la ecuación de cambio y derivando respecto al tiempo, se obtiene la relación entre las tasas de crecimiento de sus cuatro variables:

$$g_M + g_V = g_P + g_Y \tag{2}$$

donde $g_X \equiv \dot{X}/X$ es la tasa de crecimiento de la variable X. La ecuación (2) permite estudiar el efecto del crecimiento de la masa monetaria sobre la inflación y sobre el crecimiento de la actividad económica. A corto plazo el dinero no es neutral y la política monetaria influye en el empleo y la renta, y no solo en los precios. Qué parte del impulso monetario acaba en precios y qué parte en renta real no puede fijarse, sin embargo, *a priori*. Determinarlo sería la tarea de la llamada *ecuación perdida del monetarismo*.

A largo plazo, en cambio, la ecuación (2) se simplifica, dado que $g_V \approx 0$ (la velocidad del dinero se mantiene estable) y $g_Y = \bar{g}_Y$ (la renta crece a su tasa tendencial, determinada por factores reales e independiente del dinero):

$$g_P = g_M - \bar{g}_Y \;\rightarrow\; \Delta g_P = \Delta g_M$$

Es decir, a largo plazo, cualquier aumento en la tasa de crecimiento de la cantidad de dinero (agregados monetarios) conlleva un aumento equivalente en la tasa de inflación, sin efecto sobre el crecimiento del *output* real (superneutralidad del dinero respecto del *output*). El dinero es, por tanto, neutral a largo plazo, puesto que el *output* real lo fijan factores reales y la demanda de saldos reales, tal como la especifica la ecuación (1), depende de forma estable de la renta permanente, sometida igualmente a factores reales.

Estática comparativa: queda por explicar cómo se traducen las variaciones de la oferta monetaria en variaciones del *output* real a corto plazo. La respuesta está en el mecanismo de transmisión monetarista, que se deduce de las relaciones anteriores. El mecanismo se refiere a la relación directa entre el mercado de bienes y servicios y el mercado monetario que surge al incluir los saldos monetarios (reales) en la demanda de consumo y en la de dinero[21].

▪ *Consumo*: de la hipótesis de la renta permanente se deriva que el consumo (real) depende de la riqueza real de los individuos a lo largo de su vida ($C = k_P Y^P$), y no solo de su renta real. Ello abre un canal riqueza en la transmisión de ciclos a través del consumo de los hogares. A través de este canal aparecen dos tipos de efectos riqueza:

- *Efecto riqueza directo*: derivado de cambios en la cantidad de dinero y deuda pública que tienen los individuos. Formalmente, $\partial C/\partial W \cdot \partial W/\partial M^s$ o $\partial C/\partial W \cdot \partial W/\partial B$. Este tipo de efectos provocan desplazamientos adicionales de la curva IS.

- *Efecto riqueza indirecto*: derivado de cambios en el valor real de los activos que componen la riqueza, inducidos por otras variables. Este efecto opera por dos vías:

    - *Vía variaciones en el tipo de interés*: con bonos perpetuos, cualquier movimiento del tipo de interés (real) altera su valor y, a través de la riqueza, el consumo (real). Este canal hace la curva IS más elástica al tipo de interés.
    - *Vía variaciones en los precios*: la variación del nivel de precios altera el valor real del dinero y de la deuda pública en manos de los individuos (efecto Pigou). De nuevo, se producirán desplazamientos adicionales en la curva IS.

▪ *Dinero*: como en las carteras mixtas de Tobin, y con más razón al ser el dinero un bien más, dinero y bonos no son sustitutivos perfectos. Así, la emisión de deuda pública para financiar el déficit puede provocar un efecto reajuste de cartera que desplace la curva LM hacia la izquierda.

De la microfundamentación anterior se deduce la efectividad relativa de las políticas de demanda:

▪ *Política monetaria*: en el plano renta-tipo de interés (figura A5.5), un aumento de la oferta monetaria desplaza la LM a la derecha, de LM_0 a LM_1, y la economía pasa del punto 0 al 1, con más renta y un tipo de interés más bajo. Para los monetaristas, la política monetaria supera con mucho a la fiscal en eficacia.

[21] Esto se entiende al comparar la ley de Walras del IS-LM con la que plantea Friedman. De acuerdo con la aportación de Keynes, los agentes solo mantienen dinero y bonos, de manera que el cumplimiento de la ley de Walras en ambos mercados permite determinar el tipo de interés de la economía: $M^s - M^d + B^s - B^d = 0$. Si en un mercado hay un exceso, en el otro habrá un exceso compensatorio y viceversa. Así, el ajuste de las carteras solo se relaciona de forma indirecta con el mercado de bienes, ya que ni el dinero ni los bonos guardan un vínculo directo con este último. Posteriormente, autores como Metzler o Tobin amplían esta condición al considerar más activos, si bien mantienen la separación de las decisiones de inversión y las de ahorro y consumo. Friedman rompe esta relación indirecta al considerar una ley de Walras del tipo: $M^s - M^d + B^s - B^d + (Y^s - Y^d) = 0$. En consecuencia, se puede dar un exceso en el mercado monetario ($M^s - M^d$) sin que necesariamente se produzca un exceso compensatorio en el de bonos, pues puede producirse en el sector real.

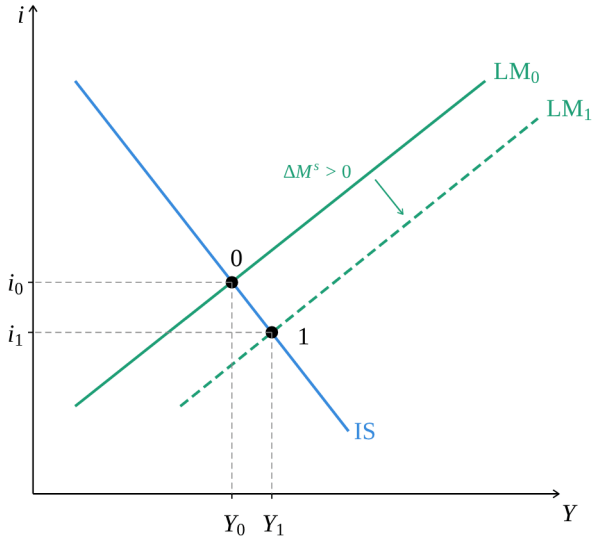

Figura A5.5: Política monetaria expansiva en el IS-LM. La LM se desplaza a la derecha, la renta aumenta y el tipo de interés desciende.

- *Política fiscal*: es muy poco efectiva:

  - *IS elástica*: un consumo guiado por la renta permanente y una inversión que responde con fuerza al tipo de interés hacen que pequeñas variaciones del tipo (real) muevan mucho la renta de equilibrio. Por eso, en el plano (Y, r), la IS es casi horizontal.

  - *LM rígida*: el dinero, como bien más, tiene muchos sustitutivos y su demanda se apoya en la renta permanente, así que la LM es muy rígida al tipo de interés (real). En el plano (Y, r), la LM es casi vertical.

Por ello, en el plano renta-tipo de interés (figura A5.6), un aumento del gasto público desplaza la IS de IS_0 a IS_1. El efecto multiplicador de la política fiscal (0-0′) es menor que en la SNC, pues el consumo depende de la renta permanente, y el efecto expulsión (0′-1) sobre la inversión y el consumo privados es elevado. Si, además, la demanda de dinero es insensible al tipo de interés ($h = 0$), la LM es vertical y la expulsión es total, es decir, el aumento del gasto público desplaza uno a uno al gasto privado sin efecto alguno sobre la renta.

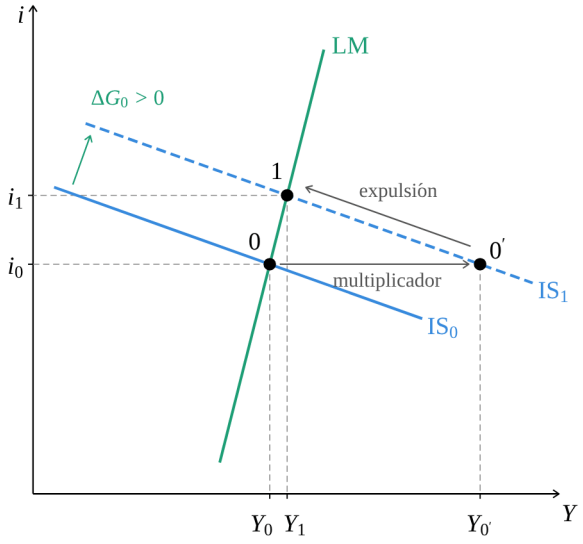

Figura A5.6: Política fiscal expansiva en el IS-LM con supuestos monetaristas (IS muy elástica y LM muy rígida). A tipo de interés constante la renta pasaría de 0 a 0′ (efecto multiplicador), pero la subida del tipo de interés expulsa gasto privado (0′-1), de modo que la renta apenas aumenta.

▪ *Pendientes*: con los multiplicadores del apartado 1.2.1, una IS plana (v elevado) y una LM casi vertical ($h \approx 0$) acercan el multiplicador fiscal a cero y el monetario a $1/k$, justo el caso opuesto al de la síntesis neoclásica (anexo).

Extensiones:

▪ *Gran Depresión*: entre 1929 y 1933, la cantidad de dinero llegó a caer más de un tercio. Para M. Friedman y A. Schwartz, en su *A Monetary History of the United States, 1867-1960* (1963), ese desplome de la liquidez fue el origen principal de la depresión, y la Reserva Federal fue su principal responsable. No frenó las oleadas de quiebras bancarias que lo provocaron e incluso endureció la política monetaria en octubre de 1931, al subir el tipo de descuento tras el abandono del patrón oro por el Reino Unido[22]. El proteccionismo generalizado agravó la crisis al hundir el comercio mundial, pero la tesis de ambos autores, estrictamente monetaria, deja fuera el proteccionismo.

▪ *Diseño e instrumentación*: los monetaristas fijan su objetivo en la cantidad de dinero, no en el tipo de interés, y como instrumento la controlan mediante operaciones de mercado abierto.

▪ *Desinflación*: Friedman recomendó con carácter general una reducción gradual y preanunciada de la tasa de crecimiento de la cantidad de dinero, que diera tiempo a los agentes a revisar sus expectativas y limitara el coste en producción y empleo. Admitía, no obstante, que la transición implicaría un periodo de menor actividad y mayor desempleo. Ante inflaciones muy elevadas, en cambio, llegó a aconsejar un "tratamiento de choque" (así lo hizo en Chile, en 1975). Lo que se buscaba, en última instancia, era una inflación baja y estable sobre la que asentar un crecimiento duradero.

[22] En 2002, en el homenaje a Friedman por su 90.º cumpleaños, Ben Bernanke, entonces miembro de la Junta de Gobernadores de la Reserva Federal, lo reconoció en nombre de la institución: "Regarding the Great Depression. You're right, we did it. We're very sorry".

▪ *El problema de los retardos*: los monetaristas desaconsejan, en general, la estabilización activa, por dos motivos:

- *Retardos*: las medidas tardan en surtir efecto, con plazos largos y variables, sobre todo en la política fiscal pero también en la monetaria, de modo que una intervención activa puede acabar siendo procíclica. De hecho, Friedman (1960) estimó que los máximos de la tasa de crecimiento del dinero se adelantaban en promedio unos 16 meses a los del ciclo y los mínimos unos 12 meses. Los adelantos iban de 6 a 29 meses en los máximos y de 4 a 22 en los mínimos.

- *Teoría aceleracionista*: un paro por debajo de su tasa natural solo se mantiene con una inflación creciente.

▪ *Política reglada*: los monetaristas quieren una política monetaria a salvo de presiones políticas, precisamente porque la consideran muy poderosa. La autoridad monetaria debe, pues, atenerse a reglas, tener la inflación como meta y renunciar a la discrecionalidad, por dos razones:

- *Corto plazo*: conviene ahorrar a la economía sacudidas monetarias, porque nadie puede anticipar qué parte de su efecto irá a precios y qué parte a renta real (ecuación perdida del monetarismo).

- *Largo plazo*: existe una tasa natural de paro y la economía tiende al pleno empleo, de modo que la política monetaria solo afecta al nivel de precios, que es lo único que puede estabilizar.

▪ *Reglas*: así, Friedman propone:

- *Regla de política monetaria* (Friedman, 1960): la tasa de crecimiento de la masa monetaria debe ser constante y ajustada a las necesidades de liquidez de la economía (regla del k %).

- *Regla de inflación óptima* (Friedman, 1969): el óptimo de Pareto exige igualar el coste privado de mantener dinero, el tipo de interés nominal, a su coste social de producción, prácticamente nulo en el dinero fiduciario. La autoridad monetaria puede reducir ese coste a cero con una deflación igual al tipo de interés real:

$$\pi = i - r \ \rightarrow \ \pi = 0 - r \ \rightarrow \ \pi = -r$$

2.2.5. La ecuación perdida del monetarismo y evidencia empírica

La teoría cuantitativa determina la renta nominal, pero no cómo se reparte entre precios y renta real, indeterminación que se conoce como la **ecuación perdida del monetarismo**. Friedman la suple con la hipótesis de la tasa natural de paro y las expectativas adaptativas, que, con la ley de Okun, dan una oferta agregada ampliada por las expectativas (apartado 2.3.2). Así, a corto plazo el reparto depende del error de previsión de los agentes, y a largo plazo la producción retorna a su nivel natural y todo el efecto recae sobre los precios.

La evidencia empírica en que se apoyó el monetarismo procede, sobre todo, del análisis histórico de Friedman y Schwartz (1963), según el cual los impulsos monetarios nacen en el propio sector monetario y van seguidos de ciclos reales, del contraste de Friedman y Meiselman (1963), que halló la velocidad del dinero más estable que el multiplicador de la inversión, y de la ecuación de St. Louis (Andersen y Jordan, 1968), que halla un efecto del dinero sobre la renta nominal mayor, más rápido y más predecible que el de la política fiscal. Sin embargo, sus críticos objetaron que la correlación entre dinero y renta podía reflejar una causalidad inversa.

Conclusiones:

▪ *Aspectos positivos*: el monetarismo devolvió al dinero un papel central en la explicación de la inflación y del ciclo, con el respaldo de una amplia evidencia histórica. Además, la defensa monetarista de las reglas frente a la discrecionalidad contribuyó a que la estabilidad de precios se convirtiera en el objetivo principal de la política monetaria.

▪ *Limitaciones*: a partir de los años ochenta, la innovación financiera y la desregulación quebraron la estabilidad de la velocidad de circulación, sobre la que descansaba la regla de crecimiento constante. Por ello, los bancos centrales abandonaron los agregados monetarios como objetivo intermedio en favor de los tipos de interés y, más tarde, de los objetivos directos de inflación.

2.3. La crítica por el lado de la oferta

Por el lado de la oferta, la controversia giraba en torno a la curva de Phillips (Phillips, 1958). Anticipándose a la inestabilidad que la curva mostraría durante los años setenta, Milton Friedman plantea en *The Role of Monetary Policy* (1968) que existe una **tasa natural de paro**, la única compatible con una inflación estable (no aceleradora)[23]. La tesis da un gran peso al lado de la oferta y al funcionamiento estructural del mercado de trabajo, e implica una curva de Phillips vertical a largo plazo en el plano (u, π), en la tasa natural.

Una inflación persistente tiene, por tanto, un origen de demanda (monetario) y nunca de costes. Friedman lo resumió al afirmar que la inflación es "siempre y en todas partes un fenómeno monetario", es decir, que solo surge cuando el dinero crece más deprisa que la producción. La frase aparece por primera vez en *Inflation: Causes and Consequences* (1963), y Friedman la repitió en *The Counter-Revolution in Monetary Theory* (1970).

2.3.1. Inestabilidad de la curva de Phillips

Para Friedman, la curva de Phillips carece de estabilidad porque su formulación original deja fuera las expectativas, al suponer ilusión monetaria y expectativas estáticas sobre los precios. Lo que a los agentes les importa, al final, es lo que su salario permite comprar, así que, al pactar una cifra nominal, la ajustan a los precios que prevén.

Friedman modela esas expectativas con la HEA. Cada periodo, la inflación esperada incorpora una parte $\lambda \in (0, 1]$ del último error de previsión, $\pi^e_{t+1} = \pi^e_t + \lambda(\pi_t - \pi^e_t)$, y ese parámetro mide lo rápido que aprenden los agentes. Al incorporar las expectativas a la relación de Phillips, la curva se desplaza conforme se alteran las expectativas.

Friedman consideraba, además, que las variables nominales (en particular, la tasa de inflación) no podían alterar permanentemente el comportamiento de las variables reales (como el desempleo), porque la ilusión monetaria podía ser plausible, pero no para siempre.

2.3.2. La curva de Phillips ampliada por las expectativas y la tasa natural de paro (Friedman, 1968; Phelps, 1967)

Origen: esta idea llevó a Milton Friedman (1968) y, de forma independiente, a Edmund Phelps (1967) a formular el concepto de tasa natural de paro. La tasa natural es la tasa de desempleo a la que finalmente llega la economía por el comportamiento optimizador de los agentes. Se trata,

[23] Friedman eligió el nombre por analogía con el tipo de interés natural de Wicksell. La definió como el paro que resultaría del sistema walrasiano de ecuaciones de equilibrio general si se incorporaran en él las características estructurales de los mercados de trabajo y de bienes, como sus imperfecciones, la variabilidad aleatoria de la oferta y la demanda, el coste de informarse sobre vacantes y trabajadores disponibles y los costes de movilidad.

por tanto, de la tasa de desempleo de equilibrio a largo plazo, que incluye el paro friccional y el estructural y está determinada por factores reales.

Idea: la inflación observada depende de la inflación esperada y del desempleo cíclico $(u_t - \bar{u})$, de modo que la disyuntiva entre inflación y desempleo solo existe mientras las expectativas no se cumplen. La relación es la *curva de Phillips ampliada por las expectativas*[24]:

$$\pi_t = \pi_t^e - \kappa_u(u_t - \bar{u}), \quad \kappa_u > 0$$

donde $\kappa_u > 0$ refleja la sensibilidad de la inflación al desempleo cíclico, $u_t - \bar{u}$, y $\bar{u}$ la tasa natural de paro. Su análoga en el modelo de salarios y precios, la tasa de paro que no acelera la inflación (NAIRU), crece con el margen de las empresas y con los factores institucionales que presionan al alza los salarios (tema A28, apartado 2.2).

> La curva de Phillips ampliada por las expectativas, con la HEA y los casos en que las expectativas se cumplen o no, se desarrolla en el tema A28, apartado 1.3.1, y la teoría aceleracionista de la inflación, en el apartado 1.3.2. Aquí se recoge su papel en la crítica monetarista.

Razonamiento: esta nueva modelización tenía dos consecuencias:

- *A corto plazo*: la inflación esperada hacía que existieran varias curvas de Phillips, cada una asociada a una expectativa diferente.

- *A largo plazo*: las expectativas acababan por recoger la inflación efectiva, de modo que precios y salarios crecían al ritmo esperado. La curva de Phillips se hacía entonces vertical y el desempleo quedaba en su nivel natural.

A partir de estas aportaciones, Friedman introdujo la *teoría aceleracionista de la inflación* (1968), que cuestionaba seriamente la concepción de la curva de Phillips original como un menú de política económica que permitía un "ajuste fino".

Desarrollo gráfico: la teoría se ilustra en el plano tasa de desempleo-inflación (figura A5.7):

1. *Curvas de Phillips a corto plazo*, $\text{CP}_{\text{CP},0}$, $\text{CP}_{\text{CP},1}$ y $\text{CP}_{\text{CP},2}$: cada una corresponde a una inflación esperada, del 0 %, el 5 % y el 10 %. Tienen pendiente negativa debido a la relación inversa entre π_t y u_t.

2. *Curva de Phillips a largo plazo*, CP_{LP}: vertical en la tasa natural de paro, $\bar{u}$, donde la inflación observada iguala a la esperada.

3. *Equilibrio*: la economía parte del punto 0, con expectativas de inflación nulas, sobre $\text{CP}_{\text{CP},0}$ y en la tasa natural de paro.

[24] Equivalentemente, $u_t = \bar{u} - \frac{1}{\kappa_u}(\pi_t - \pi_t^e)$.

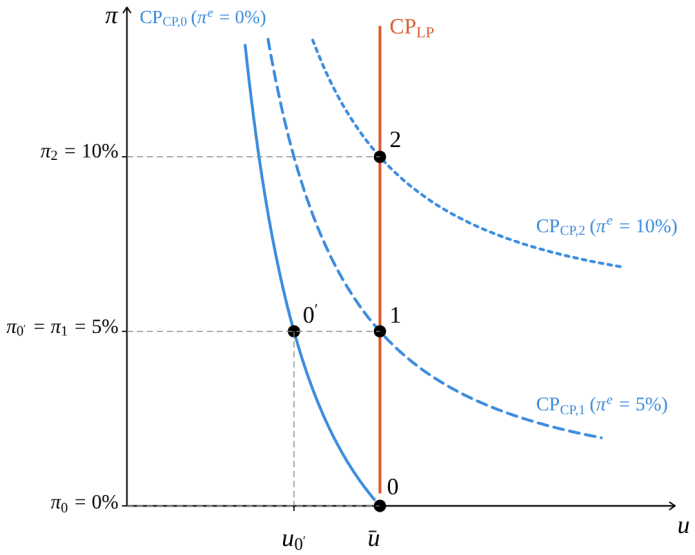

Figura A5.7: Curva de Phillips ampliada por las expectativas (Friedman, 1968; Phelps, 1967). El intercambio inflación-desempleo solo existe a corto plazo, mientras que a largo plazo la curva es vertical en la tasa natural de paro.

Dinámica: el proceso aceleracionista se desarrolla en tres etapas:

1. *Estímulo de demanda*: el Gobierno estimula la demanda agregada para reducir el paro y llevar la economía a $0'$.

2. *Error de expectativas*: en primera instancia, los agentes no revisan sus expectativas (padecen una ilusión monetaria transitoria) y, efectivamente, la tasa de desempleo se sitúa en $u_{0'} < \bar{u}$, en el punto $0'$, debido al estímulo de demanda agregada.

3. *Revisión de expectativas*: no obstante, los agentes corregirán sus previsiones un periodo más tarde. Se darán cuenta de que su salario real ha empeorado y quienes entraron en el mercado de trabajo creyendo que su salario había aumentado saldrán de él.

Así, la revisión de expectativas desplazará la curva de Phillips hacia arriba (a $\text{CP}_{\text{CP},1}$) y el aumento del desempleo llevará la economía al punto 1. El Gobierno intentará entonces acelerar la tasa de crecimiento de la demanda agregada, de modo que el proceso se repetirá (punto 2) y se desatará una espiral inflacionista capaz, en el extremo, de privar al dinero de su función de medio de cambio. Con todo, el intento será infructuoso a largo plazo, dado que, cuando los agentes acaben anticipando también la aceleración, la economía volverá a su tasa natural de desempleo.

Extensiones:

- *Política monetaria*: si se quiere que el desempleo permanezca bajo su tasa natural, los agentes deben seguir equivocándose en sus previsiones. Eso solo se consigue si la oferta monetaria crece cada periodo más deprisa que en el anterior.

- *Oferta agregada y cierre del modelo*: como en la síntesis, la ley de Okun en brechas (apartado 1.3.2) traslada la curva al plano producción-inflación. Al sustituir $u_t - \bar{u} = -(y_t - \bar{y}_t)/\phi$ se obtiene la oferta agregada ampliada por las expectativas:

$$\pi_t = \pi_t^e + \kappa\,(y_t - \bar{y}_t), \qquad \kappa = \kappa_u/\phi > 0$$

A diferencia de la de la síntesis, esta oferta tiene pendiente positiva solo a corto plazo y se desplaza al alza cada vez que sube la inflación esperada. A largo plazo, cuando $\pi_t = \pi_t^e$, es vertical en la producción potencial. Junto con la teoría cuantitativa en tasas de crecimiento, que fija el crecimiento de la renta nominal (ecuación (2)), y con la HEA, esta oferta reparte ese crecimiento entre inflación y crecimiento real. Aporta así la ecuación perdida del monetarismo (apartado 2.2.5).

- *Políticas de oferta*: a largo plazo, solo ellas podrían bajar el desempleo. Entre estas políticas figuran las que corrigen la información asimétrica en el mercado, estimulan la innovación tecnológica, forman capital humano o ajustan la fiscalidad.

Conclusiones:

- *Aspectos positivos*: explica, antes de que la estanflación lo hiciera evidente, por qué la disyuntiva entre inflación y desempleo solo es transitoria, y separa los determinantes reales del desempleo, resumidos en la tasa natural, de los monetarios de la inflación. Asimismo, el concepto de tasa natural sigue siendo la referencia para medir el paro cíclico.

- *Limitaciones*: el resultado de corto plazo descansa en expectativas adaptativas, que permiten a la autoridad engañar sistemáticamente a los agentes. En cambio, con expectativas racionales, una política monetaria anticipada no tiene efectos reales ni siquiera a corto plazo (tema A06, apartado 3.1.1). Además, la tasa natural no es observable y su estimación es imprecisa (tema A28, apartado 3.3).

2.4. Conclusiones

Aspectos positivos: hoy el monetarismo apenas cuenta en el debate económico, pero tiene el mérito de haber devuelto al centro del análisis tanto la oferta agregada, en especial el mercado laboral, como el mercado monetario.

La aportación monetarista es especialmente relevante por tres motivos:

- *Tasa natural de paro*: el concepto ha permanecido en la literatura económica como guía para medir el paro cíclico.

- *Expectativas*: el monetarismo las puso en primer plano, y las escuelas posteriores las convirtieron en un elemento central de todo modelo macroeconómico.

- *Retardos*: la insistencia del monetarismo en los de la política económica los ha hecho parte habitual de muchos modelos macroeconómicos actuales.

Limitaciones: son tanto empíricas como teóricas:

- *Recomendaciones de política económica*: en la práctica, reciben las principales críticas:

- *Medidas adoptadas*: aplicar las recetas monetaristas tuvo un coste muy alto en renta y empleo. Así lo comprobaron, en el tramo final de los años setenta y durante una parte de los ochenta, EE. UU., R. U. y otros países europeos y anglosajones, que adoptaron buena parte de ellas. El monetarismo alcanzó su máxima influencia práctica con los experimentos de control de agregados monetarios de la Reserva Federal (1979-1982) y del Banco de Inglaterra. Ambos experimentos lograron la desinflación, pero con una volatilidad de los tipos de interés muy superior a la prevista. En EE. UU., la recesión de 1981-1982 llevó el paro al 10,8 % a finales de 1982, la tasa más alta desde la Segunda Guerra Mundial hasta entonces, aunque no conviene olvidar el telón de fondo de las crisis del petróleo.

- *Objetivo intermedio*: según el modelo de Poole (1970; tema A36, apartado 2.2.1), si predominan los *shocks* monetarios (una demanda de dinero inestable), fijar como objetivo el tipo de interés estabiliza mejor la economía real que fijar el agregado. Con una demanda de dinero inestable, convertir la cantidad de dinero en objetivo intermedio acarrea serias desventajas añadidas.

- *Regla de crecimiento constante*: tampoco prosperó en la práctica la regla del k %, que pedía un ritmo fijo de crecimiento monetario. Para conducir la política monetaria, los bancos centrales optaron, en cambio, por reglas un poco más sofisticadas, como la de Taylor (1993).

▪ *Plano teórico*: también aquí la aportación monetarista presenta limitaciones:

- *HRP*: entre las muchas críticas que ha recibido destacan la falta de observabilidad de la renta permanente y el exceso de sensibilidad del consumo a la renta corriente (Flavin, 1981; tema A33, apartado 3.1.1). Aun así, su núcleo, compartido con la HCV, sigue siendo la base de la teoría moderna del consumo.

- *Ciclos*: Tobin (1970) advirtió que la correlación entre oferta monetaria y renta nominal podía deberse a una causalidad en sentido contrario. Si la renta y los precios cambian por motivos reales, la oferta monetaria puede adaptarse a ellos (dinero endógeno, vía demanda de crédito bancario y acomodación del banco central), en lugar de ser su causa.

- *Expectativas*: que las autoridades pudieran confundir una y otra vez a los agentes, como permite la HEA, resultaba inadmisible para Robert Lucas y otros autores. Esta objeción, la de más peso, surgió dentro del propio pensamiento neoclásico (tema A06).

3. Otros desarrollos posteriores

La controversia entre la síntesis neoclásica y el monetarismo dio paso a tres desarrollos posteriores, que se exponen en sus temas propios.

3.1. La nueva macroeconomía clásica

La crítica de Lucas a las expectativas adaptativas, con la que se cierra el apartado 2, enlaza con la nueva macroeconomía clásica (NMC), que sustituye la HEA por la hipótesis de expectativas racionales y explica el ciclo con mercados que se vacían continuamente.

> La nueva macroeconomía clásica se desarrolla en el tema A06, y la crítica de Lucas, en su apartado 1.3.

3.2. La nueva economía keynesiana de primera generación

La nueva economía keynesiana de primera generación (NEK1G) responde a la NMC microfundamentando las rigideces reales y nominales que la síntesis neoclásica postulaba sin justificar.

> Las teorías de las rigideces de la NEK1G se desarrollan en el tema A07, apartados 1.2 y 1.3, y el modelo de la NAIRU de Layard y Nickell, en el tema A28, apartado 2.2.

3.3. La nueva economía keynesiana de segunda generación

La nueva economía keynesiana de segunda generación (NEK2G) completa la secuencia al combinar las rigideces nominales con el método de equilibrio general dinámico de la NMC, en lo que se conoce como *nueva síntesis neoclásica*.

> El modelo de la NEK2G se desarrolla en el tema A07, apartado 2.1, y la nueva síntesis neoclásica, en el tema A06, apartado 3.4.

Conclusiones

El tema recorre el desplazamiento del centro de gravedad de la macroeconomía en la segunda mitad del s. XX. La síntesis neoclásica formalizó a Keynes en el modelo IS-LM, completó el modelo con la curva de Phillips y dominó la disciplina durante tres décadas con un supuesto dual, precios rígidos a corto plazo y flexibles a largo, que no llegó a microfundamentar. El monetarismo la criticó en su propio terreno. Por el lado de la demanda, la renta permanente y una demanda de dinero estable hacen la IS elástica y la LM rígida, de modo que la política monetaria domina a la fiscal, aunque debe ser reglada. Por el lado de la oferta, si los agentes forman expectativas, la disyuntiva entre inflación y desempleo solo es transitoria y la discrecionalidad monetaria acaba generando inflación sin ganancia real duradera.

La aportación duradera de la controversia no fue tanto la victoria de una escuela como el desplazamiento del terreno de discusión. Desde entonces, ninguna propuesta macroeconómica se considera completa si no especifica cómo se forman las expectativas y se microfundamentan sus rigideces, exigencias que la nueva macroeconomía clásica y la nueva economía keynesiana asumen como punto de partida (temas A06 y A07). Por su parte, la renta permanente se integra en las teorías del consumo intertemporal (tema A33), la demanda de dinero como activo en las de la demanda de saldos reales (tema A35) y la tasa natural en las del desempleo de equilibrio (tema A28). Por último, la disputa entre reglas y discrecionalidad sigue abierta en el diseño de la política monetaria (tema A36).

La inflación pospandemia ha devuelto al primer plano el papel del dinero. Los bancos centrales habían abandonado los agregados monetarios, y en los modelos neokeynesianos el dinero carece de papel propio (Woodford, 2008). Sin embargo, Borio, Hofmann y Zakrajšek (2023) muestran que la relación entre crecimiento monetario e inflación, casi nula cuando la inflación es baja, se acerca a la proporcionalidad cuando es alta, y que el crecimiento del dinero habría mejorado las previsiones de la inflación pospandemia. Sigue abierto, además, el debate sobre la curva de Phillips. Con las expectativas ancladas, la curva volvió a parecerse a la de los años sesenta, más plana y con una relación entre el nivel de inflación y el desempleo (Blanchard, 2016). El cambio sugiere que el alcance de la crítica de Friedman y Phelps depende del régimen de política monetaria.

Anexo. Síntesis por parámetros

Tabla A5.1: Debate sobre los parámetros del modelo IS-LM entre keynesianos y monetaristas.

Resumen del debate sobre los parámetros IS-LM	Parámetro	Keynesianos	Monetaristas
Sensibilidad de la inversión al tipo de interés	$v > 0$	Baja. La inversión está dominada por los *animal spirits*.	Alta. La inversión está dominada por la rentabilidad.
Sensibilidad de la demanda de dinero al tipo de interés[25]	$h \geq 0$	Alta. La demanda de dinero está dominada por la especulación. Solo los bonos son sustitutivos del dinero: pocas posibilidades de diversificación. Todo lleva a una alta dependencia del tipo de interés.	Baja. Múltiples activos son sustitutivos del dinero: muchas posibilidades de diversificación.
Sensibilidad de la demanda de dinero a la renta	$k > 0$	Baja. La demanda de dinero depende sobre todo del tipo de interés.	Alta. La demanda de dinero está dominada por la renta permanente.

[25] Con $h = 0$ se recupera el supuesto clásico. El dinero solo se demanda para hacer transacciones, en una proporción fija de la renta, así que M^d es insensible a cualquier movimiento del tipo de interés.

A6. La nueva macroeconomía clásica

La nueva macroeconomía clásica. La hipótesis de las expectativas racionales; la crítica de Lucas; el surgimiento de los modelos dinámicos estocásticos de equilibrio general.

Introducción

> **Introducción común.** Los temas A02 a A07 comparten una introducción, que se desarrolla al comienzo del tema **A02**. Recorre la evolución del pensamiento económico, del mercantilismo y la fisiocracia a la escuela clásica, la revolución marginalista, Keynes y la síntesis neoclásica, el monetarismo, la nueva macroeconomía clásica, la nueva economía keynesiana y la nueva síntesis neoclásica.

La estanflación de los años setenta, con inflación y desempleo altos a la vez, y el fracaso de los modelos macroeconométricos keynesianos para anticiparla abrieron el camino a la **nueva macroeconomía clásica** (NMC). La escuela radicalizó la crítica monetarista a la síntesis neoclásica (tema A05). Friedman admitía efectos reales transitorios de la política monetaria porque los agentes forman sus expectativas de forma adaptativa, mientras que la NMC los limitó a las perturbaciones que los agentes no anticipan.

Robert Lucas, Thomas Sargent, Neil Wallace y Robert Barro llevaron a la macroeconomía las **expectativas racionales** de Muth (1961) y la exigencia de **microfundamentos**, según la cual las relaciones agregadas deben derivarse del comportamiento optimizador de agentes que miran al futuro y vacían los mercados de forma continua. En los modelos de Lucas (1972), los agentes confunden las variaciones del nivel general de precios con cambios de precios relativos, y solo esa confusión da efectos reales a las perturbaciones monetarias. La **crítica de Lucas** (1976) dio a esa exigencia su forma más conocida, al mostrar que los parámetros de los modelos estimados cambian cuando cambia la política económica y que esos modelos no sirven, por tanto, para evaluarla.

En política económica, si los agentes anticipan las medidas sistemáticas, la política monetaria anticipada no altera la producción (Sargent y Wallace, 1975), la discrecionalidad genera un sesgo inflacionista que corrigen las reglas creíbles (Kydland y Prescott, 1977), y Barro (1974) devolvió actualidad a la equivalencia ricardiana. La escuela recibió críticas por el vaciado continuo de los mercados y por las exigencias informativas de las expectativas racionales, pero su método, con los modelos del ciclo real de Kydland y Prescott (1982), está en la base de los modelos de equilibrio general dinámico y estocástico que hoy utilizan los bancos centrales. Lucas (1995), Kydland y Prescott (2004) y Sargent (2011) recibieron el Premio Nobel por estas aportaciones.

El tema presenta las aportaciones de la NMC en tres partes:

- *Fundamentos*: los principales rasgos de la escuela, la hipótesis de las expectativas racionales y la crítica de Lucas, que exige microfundamentar los modelos macroeconómicos.

- *Aplicaciones*: el estudio del ciclo, con la teoría de los ciclos monetarios (Lucas) y la teoría del ciclo real (Kydland-Prescott y Long-Plosser), de cuyo método surgen los modelos de equilibrio general dinámico y estocástico.

- *Implicaciones*: las de política monetaria y fiscal, las críticas recibidas y la convergencia de la escuela hacia la nueva síntesis neoclásica.

1. Fundamentos del pensamiento de la NMC

La aportación de la NMC supuso un cambio metodológico que se asienta en dos ideas, la crítica de Lucas y la microfundamentación de los modelos macroeconómicos, y en un supuesto sobre los agentes, la hipótesis de las expectativas racionales[1].

1.1. Principales rasgos

Metodología: en gran parte a raíz de la crítica de Lucas, la NMC enfatiza la necesidad de *microfundamentar* los modelos macroeconómicos, es decir, de derivar las relaciones agregadas del comportamiento optimizador de los agentes (el nexo microeconomía-macroeconomía).

Supuestos:

▪ *Entorno*: los autores de la NMC se mueven en un mundo neoclásico, con competencia perfecta en los mercados de bienes, factores y financieros, y plena flexibilidad de precios, salarios y tipo de interés. No obstante, destacan la importancia de la información imperfecta, por la que las perturbaciones nominales no anticipadas pueden tener efectos reales transitorios sin rigideces nominales (apartado 2.1).

▪ *Agentes*: junto con la metodología, es el ámbito en el que la aportación de la NMC resulta más novedosa. Destacan dos ideas:

- *Racionalidad dinámica*: al desarrollar sus modelos en un entorno dinámico, los autores de la NMC consideran que los agentes optimizan su comportamiento en el tiempo, lo que les lleva a tomar decisiones de sustitución intertemporal (p. ej. del trabajo y el ocio entre periodos, como en Lucas y Rapping, 1969)[2]. Además, los agentes tienen una conducta previsora (*forward-looking*).

- *Hipótesis de expectativas racionales (HER)* (Muth, 1961): los agentes usan de forma eficiente la información disponible y no cometen errores sistemáticos.

Ajuste: todos estos elementos llevan más allá el pensamiento clásico y garantizan el vaciado continuo de los mercados[3].

Política económica: del paradigma anterior se deduce la defensa del no intervencionismo (apartado 3).

1.2. La hipótesis de las expectativas racionales (Muth, 1961)

Según la HER[4], a la hora de formar sus expectativas, los agentes no las ajustan gradualmente, sino que utilizan la información disponible en cada momento de la mejor manera posible y no cometen errores sistemáticos[5]:

[1] Lucas y Sargent resumieron el programa de la escuela en *After Keynesian Macroeconomics* (1979), que presentaron en un congreso de la Reserva Federal de Boston en 1978 y que suele leerse como su manifiesto.

[2] La elevada elasticidad intertemporal de la oferta de trabajo es la base de la justificación del vaciado continuo, ya que las horas trabajadas fluctúan con el salario presente en relación con el esperado.

[3] Frente a la mera tendencia hacia el equilibrio de la teoría neoclásica, el vaciado continuo implica que el equilibrio se alcanza en cada instante.

[4] Muth introduce la HER en *Rational Expectations and the Theory of Price Movements* (1961), referida a los mercados agrícolas (como alternativa a las expectativas del modelo de la telaraña).

[5] La HER supone implícitamente que las variables económicas no son puramente aleatorias, sino que obedecen a un patrón de comportamiento que los agentes pueden (y necesitan) identificar para formar sus expectativas.

$$X^e_{t+i} \equiv \mathbb{E}_t[X_{t+i} \mid \Omega_t], \qquad X_{t+i} - X^e_{t+i} = \varepsilon_{t+i}, \quad \mathbb{E}_t[\varepsilon_{t+i}] = 0$$

Es decir, la expectativa subjetiva sobre la variable X para el periodo $t+i$, formada en t, es la esperanza matemática condicionada a la información disponible, Ω_t. El error de predicción es ortogonal a dicha información (de media cero y no anticipable a partir de ella). Así, la HER admite errores, pero no sistemáticos, y no debe confundirse con las expectativas adaptativas, en las que los agentes corrigen su previsión a partir de los errores pasados.

Las principales características de la HER son:

1. *Información*: no exige información completa ni que los agentes conozcan la estructura de la economía. Si la información es completa, actúan como si conocieran el modelo, mientras que, si es incompleta, utilizan de forma óptima la disponible, cuya base no pueden ampliar por sí mismos.

2. *Posibilidad de error*: la HER admite errores, y estos pueden ser grandes cuando falta mucha información o surgen perturbaciones imprevisibles. El error se reduce a medida que aumenta la información disponible y solo desaparece cuando los agentes conocen todas las ecuaciones de un modelo determinista o realizan una previsión perfecta ($X_t = X^e_t$). Lo que la HER sí exige es que el error no se refiera a aspectos deducibles con la información disponible (de lo contrario, no la estarían utilizando de forma óptima).

3. *Conocimiento de reglas*: no es preciso conocer las reglas de política futura. Los agentes pueden tener creencias sobre ellas, que no pueden diferir de las reglas reales en aspectos deducibles con la información disponible.

4. *Ley de expectativas iteradas*: la predicción hoy de la predicción futura coincide con la predicción actual de la variable, $\mathbb{E}_t[\mathbb{E}_{t+1}(X_{t+2})] = \mathbb{E}_t(X_{t+2})$. Si se esperara que la predicción fuera a cambiar en el futuro, sería porque se dispone de información no incorporada a la actual.

Además, la HER tiene una serie de exigencias sobre el término error. Concretamente, exige que este sea ortogonal a la información disponible, lo que implica dos propiedades desde el punto de vista estadístico:

1. *Esperanza nula*: al cabo de un número grande de periodos, los errores tienen media cero, $\mathbb{E}[\varepsilon_t] = 0$.

2. *No autocorrelación*: los errores son ortogonales a la información pasada y, en particular, entre sí: $\mathbb{E}[(X_t - X^e_t)(X_{t-1} - X^e_{t-1})] = 0$.

Por el contrario, la varianza constante (homocedasticidad) y la normalidad no son requisitos de la HER, sino supuestos simplificadores que a veces se añaden (p. ej. la normalidad permite la extracción de señal lineal en el modelo de Lucas).

1.3. La crítica de Lucas (1976)

Origen: Robert Lucas, en *Econometric Policy Evaluation: A Critique* (1976)[6], cuestiona la práctica, entonces habitual, de evaluar políticas simulándolas en grandes modelos macroeconométricos de inspiración keynesiana, construidos sobre el IS-LM de la síntesis neoclásica y estimados con datos históricos.

[6] La idea tenía precedentes en las relaciones autónomas de Frisch (1938) y en Marschak (1953). La ley de Goodhart (1975), según la cual una regularidad estadística tiende a romperse en cuanto se usa como objetivo de control, expresa una intuición parecida.

Idea: la estructura de esos modelos está formada por las reglas de decisión óptimas de los agentes, y estas cambian sistemáticamente cuando cambia el comportamiento de las variables relevantes para sus decisiones, entre ellas la propia política. De ello se sigue que cualquier cambio de política altera sistemáticamente la estructura del modelo.

Razonamiento: conviene, por tanto, distinguir dos tipos de parámetros:

- *Parámetros profundos o estructurales*: las preferencias (p. ej. la impaciencia o la aversión al riesgo), la tecnología y las restricciones de los agentes. Son invariantes ante las medidas de política económica.

- *Parámetros de las reglas de decisión*: los que los modelos estiman en forma reducida (p. ej. los coeficientes de una función de consumo o de una curva de Phillips). No son invariantes, pues incorporan las expectativas de los agentes sobre la política y, por tanto, varían cuando cambia el régimen de política económica.

Los modelos que tratan los coeficientes de forma reducida como si fueran estructurales son, así, inadecuados para predecir los efectos de un cambio de política[7], porque la relación estimada bajo un régimen deja de valer justo cuando se adopta otro. Lucas ilustra el argumento con la función de consumo, la de inversión y la curva de Phillips, el caso más conocido.

Extensiones: de la crítica se derivan tres consecuencias metodológicas:

- *Evaluación de reglas*: lo que puede evaluarse no son medidas aisladas, sino reglas de política estables y conocidas por los agentes, comparando el comportamiento de la economía bajo cada una. De ahí procede el énfasis posterior en las reglas y la credibilidad (apartado 3.1.2).

- *Expectativas endógenas*: para que las expectativas respondan al régimen, los agentes han de formarlas con la información disponible sobre la política, lo que consagró la hipótesis de expectativas racionales como supuesto de referencia.

- *Modelos estructurales*: solo sirven para evaluar políticas los modelos formulados en términos de parámetros profundos, que desembocaron en los modelos de equilibrio general dinámico y estocástico (DSGE, por sus siglas en inglés), objeto del apartado 2.3.

La inefectividad de la política monetaria anticipada de Sargent y Wallace (apartado 3.1.1) aplica la misma lógica. Si la regla es conocida, su parte sistemática ya está descontada en las expectativas.

Conclusiones:

- *Aspectos positivos*: la crítica de Lucas supuso un cambio en la investigación macroeconómica, a menudo descrito como una sustitución de paradigmas (*paradigm shift*, en términos de Kuhn).

- *Limitaciones*: su relevancia empírica ha resultado, sin embargo, menor de lo que se temía. Los contrastes de superexogeneidad (Favero y Hendry, 1992) comprueban si los coeficientes de una ecuación se mantienen estables cuando cambia el proceso que siguen las variables de política, y en sus aplicaciones (p. ej. a la demanda de dinero) apenas encuentran la inestabilidad que predice la crítica.

[7] *Ejemplo*: que un banco con guardias de seguridad nunca haya sido atracado no justifica despedirlos, pues la ausencia de robos depende precisamente de su presencia. Con otra política, los potenciales ladrones reconsiderarían su conducta.

1.4. Conclusiones

La NMC conserva el mundo neoclásico de competencia perfecta y precios flexibles, y su novedad está en el modo de modelizar a los agentes. Estos optimizan a lo largo del tiempo, sustituyen trabajo y ocio entre periodos y forman expectativas racionales, que admiten errores, pero no sistemáticos. Con esos supuestos los mercados se vacían de forma continua, y las perturbaciones nominales no anticipadas pueden tener efectos reales transitorios por falta de información, sin necesidad de rigideces nominales.

La crítica de Lucas convierte esos supuestos en una exigencia de método. Si las reglas de decisión de los agentes cambian con la política, los coeficientes estimados en forma reducida no sirven para evaluarla, y solo permanecen estables los parámetros profundos de preferencias, tecnología y restricciones. De ahí que se evalúen reglas conocidas y no medidas aisladas, y que los modelos se formulen con esos parámetros. La sección 2 aplica este método al ciclo, primero con sorpresas monetarias y después con perturbaciones tecnológicas.

2. Aplicaciones y desarrollos: estudio del ciclo walrasiano

Sentados los fundamentos de la escuela, sus aportaciones más características se encuentran en la modelización del **ciclo** económico. Para Lucas, un ciclo es un movimiento autocorrelacionado de las principales variables (producción, consumo, inversión, horas trabajadas, precios o tipos de interés) alrededor de una tendencia, que puede describirse mediante una ecuación con perturbaciones estocásticas. La NMC ofrece dos explicaciones del ciclo, ambas como respuesta óptima de agentes racionales a las perturbaciones en un marco de mercados que se vacían:

1. *Teorías del ciclo nominal*: los ciclos por *shocks* monetarios no anticipados (Lucas), donde el impulso es nominal y opera a través de un problema de información imperfecta.

2. *Teoría del ciclo real*: los ciclos por *shocks* reales no anticipados (la teoría del ciclo real o TCR), donde el impulso es tecnológico y no requiere fricciones informativas.

> Ambos modelos se presentan aquí en su versión esencial, como pilares del pensamiento de la NMC sobre los que descansan sus implicaciones de política económica (apartado 3). Su lectura como teorías del ciclo (análisis gráfico, dinámica y extensiones) se desarrolla en el tema A42, apartados 2.1 y 2.2, y su contraste con los hechos estilizados, en el apartado 3.3 de ese tema.

2.1. Modelo OA-DA con información imperfecta (Lucas, 1972; 1973; 1975)

Origen: Robert Lucas, en una serie de artículos durante los años setenta, elabora un modelo de oferta y demanda agregada, conocido como *modelo de las islas de Lucas*, que reelabora con matices el macromodelo neoclásico. Se apoya en dos trabajos previos:

- *Economía de la información*: el de Edmund Phelps (1968) y, sobre todo, su artículo de 1969, recogido en el libro de 1970. De este trabajo nace la *parábola de las islas*, base del *problema de extracción de señales* (trabajadores y productores dispersos en "islas" que carecen de información sobre las variables agregadas).

- *Mercado de trabajo*: el de Lucas con Leonard Rapping (1969), del que toma la importancia de la *elasticidad de sustitución intertemporal* (ESI) en las decisiones de trabajo. Así, cuando el salario presente es alto en relación con el esperado, los agentes reducen su ocio y aumentan su oferta de trabajo[8].

Idea: la diferencia fundamental con el modelo clásico está en la modelización de la oferta. La oferta agregada a corto plazo (OA_{CP}) y la de largo plazo (OA_{LP}) difieren por las percepciones transitorias erróneas sobre los precios. De ahí surge una relación positiva entre precios y producción a corto plazo, en la que la producción responde a la diferencia entre los precios observados y los esperados.

[8] En el artículo original, no obstante, todavía se considera que la formación de expectativas se realiza de forma adaptativa.

2.1.1. Supuestos

Entorno: el modelo supone un mundo neoclásico (competencia perfecta y plena flexibilidad de precios, salarios y tipo de interés), pero con información imperfecta.

La economía se compone de múltiples mercados o "islas". Cada productor observa el precio de su mercado, p_{it}, pero no el nivel general de precios, p_t. No puede distinguir, por tanto, si una variación de su precio responde a un *shock* específico de su mercado (cambio de preferencias, ante el que lo racional es alterar la producción) o a un *shock* general de demanda agregada (ante el que conviene no alterarla). Todas las variables se expresan en logaritmos.

Agentes: cada productor produce y vende un único bien en su mercado local, al precio observado, y compra bienes en otros mercados sobre los que forma un precio esperado. Forma sus expectativas según la HER, por lo que $p_t^e \equiv \mathbb{E}_{t-1}[p_t]$ es el nivel general de precios esperado con la información previa a observar el precio local. Optimiza intertemporalmente, de modo que un precio relativo alto lleva a aumentar la producción, lo que implica más demanda de trabajo por parte de las empresas y más oferta de trabajo por parte de los hogares, en virtud de la sustitución intertemporal del ocio. En cuanto a la autoridad monetaria, la cantidad de dinero nominal se desagrega en un componente previsible o sistemático y un componente no previsible:

$$m_t = \mathbb{E}_{t-1}[m_t \mid \Omega_{t-1}] + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_m^2)$$

El término ε_t recoge las sorpresas monetarias. Agrupa lo que los agentes no pueden anticipar de la política monetaria (su parte aleatoria, no sistemática) y los errores de control monetario.

2.1.2. Modelo

Planteamiento: el modelo se resume en una función de oferta agregada, derivada del problema de extracción de señales, y una función de demanda agregada, derivada de la teoría cuantitativa:

1. *Oferta agregada*: la oferta de cada mercado se compone de un componente normal o tendencial, común a todos los mercados ($\bar{y}_t$, el *output* asociado a la tasa natural de paro), y un componente cíclico que depende de los precios relativos. En este último opera la extracción de señales.

Ante un *shock* de precios, el productor pondera la información con un coeficiente $\vartheta \in [0, 1]$, el peso de los *shocks* relativos o específicos en la varianza total de los precios. Puede leerse como la probabilidad que el productor asigna a que la perturbación afecte solo a su mercado:

$$\vartheta = \frac{\sigma_i^2}{\sigma_i^2 + \sigma_p^2}$$

En una economía con inflación alta y volátil (σ_p^2 grande) los *shocks* generales son más probables ($\vartheta \to 0$) y el productor apenas reacciona a un cambio de precios, mientras que en una economía con precios estables ocurre lo contrario ($\vartheta \to 1$).

La función de oferta de cada mercado es entonces $y_{it} = \bar{y}_t + \zeta\,(p_{it} - p_t^e)$, con $\zeta \equiv \vartheta\eta > 0$, donde η es la sensibilidad (elasticidad) de la producción a los *shocks* de precios en su mercado[9]. Así, el productor solo responde a la parte del diferencial de precios que atribuye a un *shock* relativo.

[9] Con $y = l^\alpha$, $\eta = \alpha/(1-\alpha)$.

Si todos los mercados son iguales y se añade un término $\varepsilon_t^s \sim (0, \sigma_s^2)$ que recoge los *shocks* exógenos de oferta (productividad, clima, precio de los insumos), agregando las ofertas individuales (el componente específico promedia a cero) se obtiene la *oferta agregada de Lucas*[10]:

$$\text{OA: } y_t = \bar{y}_t + \zeta\,(p_t - p_t^e) + \varepsilon_t^s$$

La producción agregada depende positivamente del componente tendencial, de la *sorpresa* de precios $(p_t - p_t^e)$ y de los *shocks* de oferta. En consecuencia, salvo por los *shocks* de oferta, no hay desviación respecto del nivel natural cuando $\vartheta = 0$ (p. ej. hiperinflación), cuando $\eta = 0$ o cuando la variación de precios está completamente anticipada $(p_t = p_t^e)$.

2. *Demanda agregada*: de forma simplificada, se parte de la teoría cuantitativa del dinero, $m_t + v_t = p_t + y_t$, donde m_t es la cantidad de dinero nominal[11] y v_t una perturbación de la velocidad de circulación (de esperanza nula) que recoge los *shocks* exógenos de demanda. La demanda agregada es, así, una función de los saldos monetarios reales:

$$\text{DA: } y_t = m_t - p_t + v_t$$

Equilibrio: igualando oferta y demanda agregadas, imponiendo la HER y despejando el precio, se obtiene la producción como función de los *shocks* de la economía[12]:

$$y_t = \bar{y}_t + \frac{\zeta}{1+\zeta}(\varepsilon_t + v_t - \varepsilon_t^s) + \varepsilon_t^s \;\rightarrow\; \tilde{y}_t = \frac{\zeta}{1+\zeta}(\varepsilon_t + v_t - \varepsilon_t^s) + \varepsilon_t^s$$

Es decir, la brecha de producción (*output gap*), $\tilde{y}_t$, es una función de los *shocks* de política monetaria, ε_t, de los *shocks* sobre la velocidad de circulación del dinero, v_t, y de los *shocks* de oferta, ε_t^s.

Desarrollo gráfico: los efectos de una expansión monetaria anticipada y no anticipada en el plano renta-precios se representan en el apartado 3.1.1 (figura A6.1), y la lectura del modelo como teoría del ciclo, en el tema A42, apartado 2.1.

Estática comparativa:

$$\frac{\partial y_t}{\partial \varepsilon_t} = \frac{\partial y_t}{\partial v_t} = \frac{\zeta}{1+\zeta} > 0, \qquad \frac{\partial y_t}{\partial \varepsilon_t^s} = \frac{1}{1+\zeta} > 0$$

Lo que mueve la producción es, por tanto, la naturaleza sorpresiva de los *shocks* monetarios y de demanda, que sin sorpresa no tienen efectos reales, y esta es la lectura principal del modelo.

▪ *Política monetaria y perturbaciones de demanda*: un cambio monetario o de demanda anticipado se traslada íntegramente a los precios y es neutral. Solo su parte no anticipada (ε_t, v_t) altera transitoriamente la producción, hasta que los agentes infieren el verdadero nivel agregado de precios.

[10] Nótese que la OA de Lucas se obtiene del comportamiento optimizador intertemporal, mientras que la OA de Layard-Nickell proviene del proceso de negociación.

[11] En sentido amplio, m_t puede interpretarse como cualquier variable nominal de política que desplace la demanda agregada (en las versiones modernas, el tipo de interés fijado por el banco central).

[12] Derivación. (i) En equilibrio, OA = DA: $\bar{y}_t + \zeta(p_t - \mathbb{E}_{t-1}p_t) + \varepsilon_t^s = m_t - p_t + v_t$. (ii) Desagregando el dinero ($m_t = \mathbb{E}_{t-1}m_t + \varepsilon_t$) y tomando expectativas con la HER, todo lo anticipable se cancela, de modo que la sorpresa de precios depende únicamente de los componentes no anticipados: $p_t - \mathbb{E}_{t-1}p_t = \frac{1}{1+\zeta}(\varepsilon_t + v_t - \varepsilon_t^s)$. (iii) Sustituyendo esta expresión en la OA se obtiene la producción de equilibrio como función de los *shocks* de la economía.

▪ *Perturbación de oferta*: los *shocks* de oferta (ε_t^s), en cambio, son perturbaciones reales, no nominales, que mueven la producción directamente. La sorpresa de precios de signo contrario que provocan solo amortigua su efecto, que se reduce a $1/(1+\zeta) < 1$.

▪ *Volatilidad de la demanda agregada*: la dependencia de ϑ respecto de σ_p^2 añade una implicación empírica contrastable (Lucas, 1973). En concreto, en países con una demanda agregada más volátil, la producción responde menos a las sorpresas nominales, de modo que la disyuntiva producción-inflación es menos favorable (la curva de Phillips a corto plazo es más vertical en el plano (u, π)). No existe, pues, una disyuntiva explotable de forma sistemática por la política económica.

Extensiones: la curva de Phillips con expectativas racionales que se deriva de la OA de Lucas se recoge en el Anexo A.2 y, en el marco de la teoría del desempleo, en el tema A28.

2.1.3. Conclusiones

La parábola de las islas de Phelps, la modelización del mercado de trabajo de Lucas y Rapping y el modelo de OA-DA sentaron los cimientos del pensamiento de la NMC y de su instrumental metodológico.

Aspectos positivos:

▪ *Microfundamentación rigurosa*: la oferta agregada se deriva del comportamiento optimizador de agentes racionales, en coherencia con la crítica de Lucas, y constituye la primera explicación microfundamentada de la no neutralidad del dinero.

▪ *No neutralidad sin rigideces*: obtiene efectos reales de los *shocks* monetarios a corto plazo con precios plenamente flexibles y vaciado continuo, sin recurrir a rigideces nominales. Concilia así la evidencia de corto plazo (relación positiva precios-producción) con la hipótesis de la tasa natural y la neutralidad de largo plazo (OA_{LP} vertical en el plano (y, p)).

▪ *Predicción nítida y contrastable*: distingue dinero anticipado (neutral) de no anticipado (real), y su predicción empírica (Lucas, 1973) recibió inicialmente respaldo.

Limitaciones:

▪ *Plausibilidad de la fricción informativa*: dada la difusión rápida y barata de las estadísticas de precios, es difícil sostener que los agentes confundan de forma persistente las variaciones relativas y agregadas hasta el punto de generar ciclos.

▪ *Magnitud y persistencia*: las sorpresas monetarias explican solo una pequeña parte de la variabilidad del producto y del empleo. Sobre todo, el modelo carece de un mecanismo de propagación (el *output* no hereda sus valores pasados), por lo que no reproduce la persistencia del ciclo.

▪ *Elasticidad de la oferta de trabajo*: para que las sorpresas generen fluctuaciones de empleo apreciables se requiere una elevada elasticidad de sustitución intertemporal del trabajo, contraria a la evidencia microeconómica.

▪ *Vaciado continuo*: el supuesto excluye el desempleo involuntario, en tensión con los hechos estilizados, pues la mayor parte del ajuste del empleo se produce por flujos entre empleo y desempleo, no por las horas de los ocupados.

A estas limitaciones se sumó, a finales de los años setenta y comienzos de los ochenta, la evidencia econométrica directa. Barro (1977; 1978) halló que solo el crecimiento monetario no anticipado afectaba al desempleo y a la producción. Sus estimaciones, sin embargo, atribuían a esas sorpresas efectos que se prolongaban unos dos años, difíciles de explicar por una confusión informativa cuando en EE. UU. los agregados monetarios se publicaban cada semana. Mishkin (1982) encontró, además, que el dinero anticipado también tenía efectos reales, no menores que los del no anticipado, con lo que el modelo perdía su predicción distintiva. La respuesta dentro de la propia escuela fue conservar el método y cambiar el impulso.

2.2. Ciclos debidos a *shocks* reales no anticipados (Kydland y Prescott, 1982; Long y Plosser, 1983)

Origen: Kydland y Prescott (1982) y Long y Plosser (1983) sientan las bases de la teoría del ciclo real (TCR) intentando replicar la dinámica observada en la economía estadounidense entre 1950 y 1979. El modelo de Kydland y Prescott incorporaba un plazo de construcción de varios periodos (*time to build*) como mecanismo de propagación y se calibraba y resolvía numéricamente, mientras que Long y Plosser obtuvieron una versión con solución analítica.

Idea: los ciclos se generan no por un problema de extracción de señales, sino por el comportamiento optimizador de los agentes ante perturbaciones reales, de carácter tecnológico, estocásticas y autocorrelacionadas.

2.2.1. Supuestos

Entorno: la economía es cerrada, con información perfecta, tiempo discreto y horizonte infinito, competencia perfecta, mercados completos y plena flexibilidad de precios, salarios y tipo de interés[13].

Agentes: se cumple la HER y existe un número elevado de agentes idénticos (simultáneamente consumidores, productores y oferentes de trabajo), con vida infinita (altruismo intergeneracional perfecto). Su único activo es el capital, k_t, por lo que sus rentas son salariales y del capital, y sus preferencias se recogen en una función de utilidad aditivamente separable en el tiempo[14] que considera consumo y ocio. Se considera que no participa el sector público, lo que permite aislar los efectos de las perturbaciones tecnológicas.

Tecnología: la producción sigue una función Cobb-Douglas con tecnología neutral en el sentido de Harrod y rendimientos constantes a escala, afectada por perturbaciones tecnológicas autocorrelacionadas. El capital se deprecia a una tasa constante $\delta \in (0, 1]$. El caso de depreciación completa ($\delta = 1$), el de Long y Plosser, es el que permite resolver el modelo analíticamente[15].

2.2.2. Modelo

El modelo se resume en cuatro ecuaciones de comportamiento de hogares y empresas, tres ecuaciones que describen la tecnología y una condición de equilibrio competitivo, lo que permite estudiar la dinámica del *output* (el ciclo).

Hogares: el hogar representativo consume, trabaja y ahorra.

[13] Por tanto, los agentes pueden mover fondos libremente entre periodos al tipo de interés de mercado, r_t.

[14] Supuesto cómodo analíticamente pero restrictivo, porque ignora la inercia y los hábitos de consumo que revelan los estudios empíricos, así como las decisiones de formación.

[15] Romer (2012, p. 201) resuelve este caso particular, sin sector público y con $\delta = 1$.

▪ *Planteamiento*: maximiza el valor actual esperado de su utilidad a lo largo del tiempo, sujeto a su restricción presupuestaria intertemporal (RPI), en la que el capital k_t es el único activo[16]:

$$\max_{c_t, l_t, k_{t+1}} U = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(c_t, 1-l_t), \qquad \beta \in (0,1)$$
$$\text{s.a.:} \begin{cases} c_t + k_{t+1} - (1-\delta)k_t = w_t l_t + (r_t + \delta)k_t \\ k_0 \text{ dado} \end{cases}$$

donde c_t es el consumo, l_t el trabajo ofrecido (y $1 - l_t$ el ocio), k_t el capital, único activo del hogar, $u(\cdot)$ la utilidad instantánea en consumo y ocio, β el factor de descuento, k_0 el capital inicial, $\mathbb{E}_0$ la expectativa condicionada a la información del periodo inicial, w_t el salario, r_t la remuneración neta del capital y δ la tasa de depreciación. El lado izquierdo es el consumo más la inversión bruta, $i_t = k_{t+1} - (1-\delta)k_t$, y el lado derecho la renta del hogar, con el capital remunerado a su coste de uso bruto, $r_t + \delta$. La restricción equivale a $c_t + k_{t+1} = w_t l_t + (1 + r_t)k_t$.

▪ *CPO*: con solución interior y la restricción agotada, derivando el lagrangiano respecto a c_t, l_t y k_{t+1}, las CPO son la condición intratemporal entre ocio y consumo y la ecuación de Euler estocástica, a las que se suma la condición de transversalidad, que anula el valor actual esperado del capital que queda al final del horizonte. La condición intratemporal es la misma condición de oferta de trabajo que en un contexto estático, puesto que se refiere a un mismo periodo. Según la ecuación de Euler, en el óptimo la utilidad que aporta una unidad más de consumo presente coincide con la utilidad marginal esperada y descontada que esa misma unidad, ahorrada, reportaría un periodo después. En caso contrario, al agente le compensaría desplazar consumo entre periodos para elevar su utilidad[17]:

$$\text{RMS}_{c_t\,(1-l_t)} = \frac{\partial u/\partial(1-l_t)}{\partial u/\partial c_t} = w_t, \qquad \frac{\partial u}{\partial c_t} = \beta\, \mathbb{E}_t\left[\frac{\partial u}{\partial c_{t+1}}\,(1 + r_{t+1})\right]$$
$$\lim_{t\to\infty} \mathbb{E}_0\left[\beta^t \frac{\partial u}{\partial c_t}\, k_{t+1}\right] = 0$$

▪ *CSO*: la concavidad estricta de u en consumo y ocio y la de la tecnología garantizan que las CPO, con la condición de transversalidad, son necesarias y suficientes.

Empresas: producen con trabajo y capital.

▪ *Planteamiento*: maximizan su beneficio[18], $\max_{l_t, k_t} \Pi_t = y_t - w_t l_t - (r_t + \delta)k_t$.

▪ *CPO*: el trabajo se remunera al salario y el capital a su coste de uso (tipo de interés más depreciación), cada factor según su productividad marginal. A pesar del contexto dinámico, se llega a condiciones análogas a las del contexto estático:

$$\underbrace{\frac{\partial f}{\partial l_t}}_{\text{Pmg}_l} = w_t, \qquad \underbrace{\frac{\partial f}{\partial k_t}}_{\text{Pmg}_k} = \underbrace{r_t + \delta}_{\text{coste de uso}}$$

[16] Los autores de la NMC se apoyan en el modelo de Ramsey-Cass-Koopmans con dos diferencias fundamentales: la consideración del ocio en la función de utilidad y la aleatoriedad de los valores de la rentabilidad y del salario debido a la aleatoriedad de la tecnología (esto introduce la esperanza en la función).

[17] El tipo r_{t+1} no es independiente de c_{t+1} (a diferencia del modelo de Hall), ya que r_t no es constante y depende de los *shocks* tecnológicos del modelo.

[18] Se considera que el *shock* se realiza antes de que las empresas tomen sus decisiones de empleo y capital, por lo que no hace falta el operador esperanza.

▪ *CSO*: la tecnología Cobb-Douglas es cóncava en (k, l) y estrictamente cóncava en cada factor, luego las CPO caracterizan el máximo.

Tecnología: la producción es Cobb-Douglas y la tecnología crece en torno a un nivel tendencial, $\bar{A}$, a un ritmo exógeno y constante, g_A. Está afectada, además, por un componente aleatorio, a_t, que sigue un proceso AR(1) (depende de las perturbaciones previas, $\rho_A a_{t-1}$, y de un ruido blanco, ϕ_t):

$$y_t = k_t^{\alpha}(A_t l_t)^{1-\alpha}, \qquad \alpha \in (0,1)$$
$$\ln A_t = \ln \bar{A} + g_A\, t + a_t, \qquad a_t = \rho_A a_{t-1} + \phi_t, \quad \rho_A \in (-1,1)$$

Equilibrio: con vaciado de los mercados de trabajo y capital, el gasto en consumo e inversión coincide con la renta de los hogares:

$$c_t + i_t = y_t, \qquad i_t = k_{t+1} - (1-\delta)k_t$$

En el caso analíticamente resoluble ($\delta = 1$ y utilidad logarítmica, $u(c_t, 1-l_t) = \ln c_t + b\ln(1-l_t)$, donde b es el peso del ocio), se obtienen los valores solución de las variables relevantes en función de los parámetros y variables del modelo:

$$k^*_{t+1} = h_k(A_t, \alpha, k_t, l_t, \beta), \qquad c^*_t = h_c(A_t, \alpha, k_t, l_t, \beta), \qquad l^* = h_l(\alpha, \beta, b)$$

El empleo se mantiene constante, ya que solo depende de parámetros estructurales constantes y exógenos (α, β y b): $l^* = \frac{1-\alpha}{(1-\alpha)+b(1-\alpha\beta)}$. El ahorro (el capital acumulado al final de cada periodo, con una tasa de ahorro constante $s = \alpha\beta$) y el consumo, en cambio, sí se ven afectados por los *shocks* tecnológicos. La constancia del empleo no es una propiedad general de los modelos del ciclo real, sino de esta versión (utilidad logarítmica y depreciación completa).

Dinámica: a partir de los valores solución se obtienen las ecuaciones dinámicas y estocásticas de cada variable endógena[19], expresadas como desviación respecto de su trayectoria determinista[20]. En el caso del *output*, la ecuación dinámica estocástica de la desviación del *output* respecto de su tendencia determinista es:

$$\tilde{y}_t = (\alpha + \rho_A)\,\tilde{y}_{t-1} - \alpha\rho_A\,\tilde{y}_{t-2} + (1-\alpha)\,\phi_t$$

Es decir, el ciclo sigue un proceso autorregresivo de orden 2, de modo que, cuando se produce un *shock* tecnológico, el PIB se desvía de su tendencia en función de los ciclos de los dos periodos anteriores y de la persistencia de la perturbación. La duración del ciclo será mayor cuanto mayor sea la autocorrelación de la perturbación, ρ_A, que depende de la calibración del modelo[21]. Si, además, los *shocks* de productividad son frecuentes y permanentes (el caso límite de persistencia, con raíz unitaria, $\rho_A = 1$), alteran la propia tendencia (Nelson y Plosser, 1982) y explicarían así tanto el ciclo a corto plazo como el crecimiento a largo plazo.

[19] Como el objeto de estudio es el ciclo, la exposición se centra en el *output*, pero puede demostrarse que el ahorro y el consumo, en este modelo, también siguen un AR(2).

[20] Esquema de la derivación (variables en logaritmos, desviaciones respecto de la tendencia). (i) Log-linealizando la función de producción: $\ln y_t = \alpha \ln k_t + (1-\alpha)\ln A_t + (1-\alpha)\ln l_t$, con $\ln l_t = \ln l$ constante. (ii) Con depreciación total y utilidad logarítmica, la tasa de ahorro es constante, $k_{t+1} = s\, y_t$, de modo que $\ln k_t = \ln s + \ln y_{t-1}$. (iii) Sustituyendo y restando la tendencia se llega a $\tilde{y}_t = \alpha\,\tilde{y}_{t-1} + (1-\alpha)\,a_t$ y, introduciendo el proceso AR(1) de la tecnología $a_t = \rho_A a_{t-1} + \phi_t$, se obtiene el AR(2) del cuerpo del tema.

[21] *Ejemplo* (con $\alpha = 1/3$): si $\rho_A = 0$, bastan dos periodos para que desaparezca casi el 90 % del efecto inicial. Incluso con $\rho_A = 0{,}5$, dos tercios del efecto inicial desaparecen en tres periodos.

Bienestar: como la economía es walrasiana, el ciclo es la respuesta óptima de los agentes a los *shocks*. Por tanto, las fluctuaciones no constituyen un fracaso de los mercados, sino la respuesta eficiente (Pareto-óptima) de la economía a las perturbaciones, y toda intervención pública para mitigarlas reduce el bienestar.

> Los efectos de un *shock* tecnológico sobre cada variable, el análisis gráfico del ciclo y las extensiones del modelo (indivisibilidades del trabajo, depreciación incompleta, sector público) se desarrollan en el tema A42, apartado 2.2. El modelo de crecimiento sobre el que se construye la TCR, el de Ramsey-Cass-Koopmans, se estudia en el tema A43, apartado 3.

2.2.3. Conclusiones

Aspectos positivos: con modelos relativamente sencillos se reproducía gran parte de los hechos estilizados del ciclo de EE. UU.:

- *PIB*: los autores concluían que los *shocks* tecnológicos (de productividad) podían explicar en torno al 70 % de la volatilidad del PIB de posguerra en EE. UU. (Kydland y Prescott, 1991)[22]. Los modelos del ciclo real aportaron, además, un método que después adoptó también la nueva economía keynesiana (NEK).

- *Consumo e inversión*: en sus versiones calibradas (con depreciación incompleta), el modelo recoge la suavización del consumo y el exceso de volatilidad de la inversión.

Limitaciones: no obstante, el modelo presenta problemas tanto en su mecanismo de impulso y propagación como en el comportamiento de variables distintas del PIB:

- *Impulso y propagación*: el impulso descansa en *shocks* tecnológicos, cuando en el corto plazo el ritmo del progreso técnico apenas cambia. Con un impulso tan estable es difícil generar ciclos, y menos aún recesiones (¿una regresión tecnológica explica realmente una recesión?). Además, el mecanismo de propagación del modelo básico (la acumulación de capital) es débil, de modo que una perturbación tecnológica transitoria no se traduce en una variación significativa y duradera de la producción.

- *Tasa de ahorro y oferta de trabajo constantes*: en la versión básica ello implica que inversión y consumo serían igualmente volátiles, cuando en los datos de EE. UU. la inversión es unas tres veces más volátil que la producción y el consumo de bienes no duraderos lo es menos (Rebelo, 2005).

- *Mercado de trabajo*: en el modelo, el salario real acompaña al ciclo con mucha más fuerza que en los datos, donde solo lo hace de forma moderada. Además, reproducir la amplitud de las fluctuaciones agregadas exige una oferta de trabajo muy elástica, y las elasticidades de sustitución intertemporal estimadas con microdatos son un orden de magnitud menores que las necesarias (Chetty, Guren, Manoli y Weber, 2011).

- *Neutralidad del dinero*: es una implicación central de la TCR y se mantiene incluso en versiones con dinero (p. ej. modelos *cash-in-advance*). Sin embargo, aunque hay consenso teórico sobre la neutralidad a largo plazo, la evidencia empírica muestra que no siempre se cumple a corto

[22] Los *shocks* tecnológicos se miden con el residuo de Solow, la parte del crecimiento del producto que no explican el capital y el trabajo. Summers (1986), en su réplica a Prescott (1986), objetó que ese residuo recoge errores de medida y el atesoramiento de mano de obra, de modo que no puede tomarse como una perturbación exógena.

plazo (p. ej. la ecuación de St. Louis de Andersen y Jordan, 1968, o los modelos de vectores autorregresivos, VAR, de Christiano, Eichenbaum y Evans, 1999).

2.3. Comparación y legado: los modelos DSGE

Los dos modelos comparten el núcleo del pensamiento de la NMC: agentes racionales que optimizan intertemporalmente, expectativas racionales y mercados que se vacían de forma continua. Por ello, en ambos el ciclo es la respuesta óptima de los agentes a las perturbaciones y no un fallo del sistema.

Difieren en el impulso y, con ello, en el papel del dinero:

- *Lucas*: el impulso es una sorpresa nominal, y la política monetaria no anticipada tiene efectos reales transitorios.
- *TCR*: el impulso es una perturbación tecnológica, y el dinero es neutral incluso a corto plazo.

La TCR aborda, además, la principal carencia empírica del modelo de Lucas, la falta de persistencia, ya que la acumulación de capital actúa como mecanismo de propagación y la autocorrelación de los *shocks* traslada su persistencia al ciclo. La propagación interna del modelo básico es, con todo, débil.

El legado más duradero de la TCR no fue su explicación del ciclo, sino su método. De él surgen los **modelos DSGE**, modelos de equilibrio general en los que hogares y empresas optimizan intertemporalmente con expectativas racionales, la economía está sometida a perturbaciones aleatorias y el comportamiento se describe con parámetros profundos de preferencias y tecnología. Ello los hace, en principio, robustos a la crítica de Lucas (apartado 1.3). Kydland y Prescott (1982) no estimaban esos parámetros, sino que los calibraban. La calibración consiste en asignarles valores tomados de estudios microeconómicos o de promedios de largo plazo de los datos y en juzgar el modelo por su capacidad para reproducir las volatilidades y correlaciones del ciclo.

> El modelo canónico de la NEK se desarrolla en el tema A07, apartado 2.1, y como teoría del ciclo en el tema A42, apartado 2.4. La ausencia del sistema financiero en el modelo básico y la agenda posterior a la crisis se estudian en el tema A07, apartados 3.4.2 y 3.5, y el acelerador financiero, en el tema A42, apartado 2.3.2. Aquí se recoge cómo surgieron los modelos DSGE a partir de la TCR.

Desde los años noventa, el programa evolucionó en dos direcciones que acabaron fundiéndose. La primera fue la incorporación de las rigideces de la NEK. Así, la competencia monopolística y los precios *à la* Calvo dieron lugar al modelo canónico de tres ecuaciones. Christiano, Eichenbaum y Evans (2005) mostraron, además, que un modelo de tamaño medio con rigideces nominales de precios y salarios y con rigideces reales reproduce la inercia de la inflación y la persistencia del producto tras una perturbación monetaria. Las rigideces reales del modelo son los hábitos de consumo, los costes de ajuste de la inversión y la utilización variable del capital.

La segunda dirección fue el paso de la calibración a la estimación. Smets y Wouters estimaron modelos de este tipo para la zona del euro (2003) y para EE. UU. (2007) con técnicas bayesianas, que combinan la información previa sobre los parámetros con la que contienen los datos, y el de EE. UU. obtuvo previsiones fuera de la muestra comparables a las de los modelos VAR bayesianos.

Desde mediados de los años dos mil, los bancos centrales emplean modelos DSGE estimados en sus previsiones y en el análisis de políticas, junto a modelos semiestructurales y estadísticos:

- *Reserva Federal*: el FRB/EDO (*Estimated, Dynamic, Optimization-based model*), desde 2006.

- *Banco Central Europeo (BCE)*: el *New Area-Wide Model* (NAWM), ampliado con un sector financiero en el NAWM II.

- *Banco de Inglaterra*: COMPASS (*Central Organising Model for Projection Analysis and Scenario Simulation*), desde finales de 2011.

La Gran Recesión expuso los límites de los modelos DSGE, porque los modelos dominantes en su mayoría carecían de fricciones financieras, descansaban en un agente representativo y en aproximaciones lineales en torno al estado estacionario, y trataban mal el límite inferior de los tipos de interés (*zero lower bound*, ZLB). Sin embargo, la respuesta mayoritaria ha sido ampliar el marco antes que abandonarlo. En concreto, se han añadido fricciones financieras, cuyo antecedente es el acelerador financiero, y agentes heterogéneos (modelos HANK, *Heterogeneous Agent New Keynesian*, de Kaplan, Moll y Violante, 2018). En los modelos HANK, la política monetaria actúa sobre el consumo sobre todo a través de efectos indirectos sobre la renta.

2.4. Conclusiones

Los dos modelos de la sección leen el ciclo como la respuesta óptima de agentes racionales a perturbaciones en mercados que se vacían, y difieren en el origen de esas perturbaciones. En el de Lucas, el productor que no distingue si la subida de su precio es relativa o general responde a las sorpresas monetarias, de modo que solo el dinero no anticipado altera, y de forma transitoria, la producción. La falta de persistencia y la evidencia de Barro y de Mishkin llevaron a sustituir ese impulso por uno tecnológico, cuya autocorrelación, unida a la acumulación de capital, da persistencia al ciclo de la TCR.

El método de la TCR duró más que su explicación del ciclo y dio lugar a los modelos DSGE, que incorporaron después las rigideces de la NEK y la estimación bayesiana. De ese marco se derivan las implicaciones de política de la sección 3. Si el ciclo es óptimo y solo las sorpresas tienen efectos reales, la política de estabilización es, en el mejor de los casos, inútil y, en el peor, perjudicial.

3. Implicaciones del pensamiento de la NMC

Del marco anterior se derivan implicaciones de política monetaria y de política fiscal. Se exponen primero los principales resultados y, después, las críticas a la escuela y su posterior convergencia hacia la nueva síntesis.

3.1. Implicaciones de política monetaria

En materia de política monetaria, la NMC obtiene dos resultados, la inefectividad de la política anticipada y la inconsistencia dinámica de la política discrecional.

3.1.1. Inefectividad de la política monetaria (Sargent y Wallace, 1975; 1976)

Origen: el primer resultado en materia de política monetaria se debe a Sargent y Wallace (1975; 1976).

Idea: las políticas que respondan a información de conocimiento público (como los principales indicadores macroeconómicos) serán ineficaces para estabilizar la producción.

Razonamiento:

1. *Política reglada y creíble*: cuando la política monetaria sigue una regla (es fácilmente anticipable) y goza de credibilidad, no tiene efectos reales y todo su impacto recae sobre los precios: $\partial \mathbb{E}_{t-1}[p_t]/\partial \mathbb{E}_{t-1}[m_t] = 1$. La razón es que los agentes aprenden a anticipar las posibles reglas de estabilización contracíclica, siempre que los bancos centrales las utilicen de forma sistemática.

2. *Política no creíble*: sin credibilidad, incluso el componente "anticipado" puede tener efectos reales. Supóngase que la autoridad monetaria comunica una subida de m_t que eleve la inflación, p. ej. del 0 % al 5 %. Si el público desconfía del anuncio, preverá una expansión monetaria por encima o por debajo de la anunciada, y la ejecución de la política acabará afectando a las variables reales.

Con todo, la NMC no recomienda recurrir a sorpresas monetarias, porque reducen la credibilidad futura de la política monetaria y aumentan la incertidumbre.

Desarrollo gráfico: los efectos de una expansión monetaria se representan en el plano renta-precios (figura A6.1):

1. *Oferta agregada a largo plazo*, $\mathrm{OA}_{\mathrm{LP}}$: recoge la producción cuando los precios se anticipan ($p_t = p_t^e$). Es vertical en la producción natural, $\bar{Y}$, porque sin sorpresa de precios la producción no se desvía de su nivel natural.

2. *Oferta agregada a corto plazo*, $\mathrm{OA}_{\mathrm{CP},0}$: es la oferta de Lucas para unos precios esperados dados. Es creciente porque cada productor atribuye parte de la subida de su precio a un *shock* relativo y eleva su producción.

3. *Demanda agregada*, DA_0: recoge la teoría cuantitativa para una cantidad de dinero dada. Es decreciente porque una subida de los precios reduce los saldos monetarios reales.

4. *Equilibrio*: en el punto 0, donde se cortan las tres curvas, con producción $\bar{Y}$ y precios P_0 iguales a los esperados.

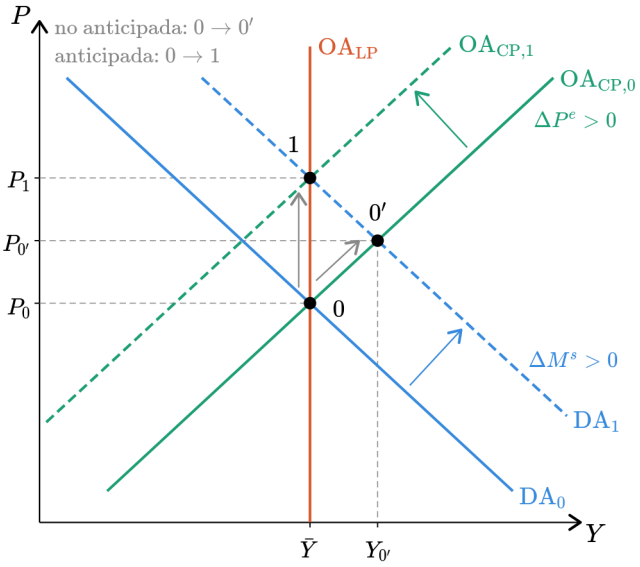

Figura A6.1: Modelo de la NMC. Con expectativas racionales solo la política monetaria no anticipada afecta a la producción. El gráfico representa las variables en niveles (P, Y, $\bar{Y}$), mientras que el texto las expresa en logaritmos.

Estática comparativa: una expansión monetaria ($\Delta M^s > 0$) desplaza la demanda agregada de DA_0 a DA_1:

- *No anticipada*: los precios esperados no cambian y la economía pasa del punto 0 al $0'$, con más producción, $Y_{0'}$, y precios $P_{0'}$.

- *Anticipada*: los precios esperados suben ($\Delta P^e > 0$), la oferta a corto plazo se desplaza a $\text{OA}_{\text{CP},1}$ y la economía pasa directamente del punto 0 al 1, con la misma producción, $\bar{Y}$, y precios P_1.

Conclusiones:

- *Aspectos positivos*: introduce la distinción entre política anticipada y no anticipada y muestra que las conclusiones de política dependen de cómo se modelicen las expectativas. Refuerza, además, el argumento a favor de reglas sencillas y predecibles frente a la política activa de estabilización. Sargent (1982) halló, por su parte, que en Austria, Hungría, Polonia y Alemania las hiperinflaciones de los años veinte terminaron con un cambio de régimen creíble, con reformas fiscales e institucionales, casi sin coste en producción.

- *Limitaciones*: el resultado exige, además de expectativas racionales, precios y salarios plenamente flexibles. Con contratos nominales plurianuales o escalonados, la política anticipada tiene efectos reales aunque las expectativas sean racionales (tema A07, apartado 1.4.1), y los contrastes empíricos no confirman que solo la política no anticipada tenga efectos reales.

3.1.2. Inconsistencia dinámica de la política monetaria (Kydland y Prescott, 1977; Barro y Gordon, 1983a)

Origen: el segundo resultado en materia de política monetaria se relaciona con la sostenibilidad de las políticas monetarias sorpresivas. Se debe a las aportaciones de Kydland y Prescott, *Rules*

Rather than Discretion: The Inconsistency of Optimal Plans (1977), en el marco de la tasa natural de paro[23], y de Barro y Gordon (1983a), en el de la teoría de juegos.

Idea: estos autores plantean si mantener en el tiempo una política monetaria sorpresiva es coherente con la HER. La respuesta se apoya en la noción de *inconsistencia dinámica*, origen del problema de la credibilidad:

1. *Teoría de juegos*: un plan es dinámicamente inconsistente si, aun sin cambiar las preferencias del decisor ni llegar nueva información, el plan que era óptimo en un momento dado deja de serlo en un periodo posterior, cuando llega el momento de ejecutarlo.

2. *Teoría económica*: una política (anunciada) es dinámicamente inconsistente si los agentes anticipan que las autoridades romperán el compromiso planteado en su anuncio, porque mantener dicha política no es óptimo a lo largo del tiempo. Por ello, la política económica (anunciada) no es creíble.

El incentivo a romper el compromiso nace de la HER, dado que los agentes económicos incorporan los anuncios de política a sus decisiones en el momento en el que se producen.

Además, la autoridad monetaria es consciente de cómo forman sus expectativas los agentes y de que puede influir en la inflación a través de sus políticas. Sabe también que existe a corto plazo una relación estadística positiva entre el *output* real (o el empleo) y la inflación. Por ello, tiene incentivos para romper sus compromisos en el corto plazo, con el fin de alcanzar diferentes objetivos cortoplacistas[24].

No obstante, los agentes, conscientes de estos incentivos, los anticipan. En consecuencia, una política monetaria *discrecional*, es decir, no basada en reglas, no logra efectos reales ni siquiera a corto plazo, porque los agentes anticipan la inflación que elegirá, y da lugar a una pérdida de bienestar (el sesgo inflacionista). Kydland y Prescott introducen así la importancia de la credibilidad.

Supuestos:

- *Entorno*: información completa, juego único (modelo estático) y mundo neoclásico, con precios y salarios flexibles, aunque los contratos nominales se fijan antes de que la autoridad decida la inflación.

- *Agentes*: intervienen dos:

  - *Sociedad*: forma sus expectativas de inflación según la HER, $\pi_t^e = \mathbb{E}_{t-1}[\pi_t]$, y sobre ellas fija sus contratos de precios y salarios.

  - *Gobierno*: intenta minimizar la pérdida social derivada de las desviaciones del desempleo y de la inflación respecto de sus objetivos. Elige un objetivo de inflación, π^T, y un objetivo de desempleo, u^T, inferior a la tasa natural (un objetivo de producción superior a su nivel natural).

Las preferencias del Gobierno reflejan las preferencias sociales y se recogen en una función de pérdida social: $L_t = (u_t - u^T)^2 + b(\pi_t - \pi^T)^2$. En ella, u_t es la tasa de paro y $u^T < \bar{u}$, el objetivo de desempleo del Gobierno (inferior a la tasa natural). El parámetro $b > 0$ refleja la importancia relativa de los dos objetivos, es decir, el grado de aversión a la inflación (el peso de la inflación en la función de pérdida en relación con el del desempleo):

[23] La evidencia empírica mostraba que las autoridades monetarias llevaban a cabo aumentos de la oferta monetaria tras incrementos en la tasa de desempleo (Barro, 1977).

[24] Ello supone cierta miopía de la autoridad. Objetivos de este tipo son elevar puntualmente la producción, reducir temporalmente el desempleo o reducir el valor real de la deuda pública mediante una inflación sorpresiva.

1. *b alto*: la autoridad monetaria pondrá más énfasis en las desviaciones de inflación que en las desviaciones del desempleo. El Gobierno es duro (*hard-nosed*)[25], como ilustra el mandato único de estabilidad de precios del BCE.

2. *b bajo*: la autoridad monetaria pondrá más énfasis en la estabilización del ciclo. El Gobierno es blando (*wet*), como ilustra el mandato dual de la Reserva Federal (Fed), que suma el máximo empleo a la estabilidad de precios. Los dos mandatos solo ejemplifican el peso de cada objetivo, porque el modelo no clasifica bancos centrales reales.

La restricción del Gobierno vendrá dada por la curva de Phillips a corto plazo, según la cual la tasa de inflación y la tasa de desempleo están inversamente relacionadas: $\pi_t = \pi_t^e - \kappa_u(u_t - \bar{u}), \quad \kappa_u > 0$. Esta relación desaparece en el largo plazo, donde $\pi_t = \pi_t^e$.

Planteamiento: la resolución del programa dará lugar a una ecuación de comportamiento discrecional de la autoridad monetaria:

$$\min_{\pi_t} L_t = (u_t - u^T)^2 + b(\pi_t - \pi^T)^2 \quad \text{s.a.: } u_t = \bar{u} - \gamma(\pi_t - \pi_t^e),\ \gamma = 1/\kappa_u$$

donde π_t es la tasa de inflación que elige la autoridad monetaria, L_t la pérdida social, u_t la tasa de paro, u^T y π^T los objetivos de desempleo e inflación del Gobierno, b el grado de aversión a la inflación, $\bar{u}$ la tasa natural de paro, π_t^e la inflación esperada por la sociedad y $\gamma = 1/\kappa_u$ la inversa del valor absoluto de la pendiente de la curva de Phillips en el plano (u, π).

CPO: la condición iguala el coste marginal de desviar la inflación de su objetivo, $2b(\pi_t - \pi^T)$, con el beneficio marginal de reducir el desempleo por debajo de la tasa natural a lo largo de la curva de Phillips. Como el Gobierno persigue un desempleo inferior al natural ($u^T < \bar{u}$), le compensa generar inflación por encima de π^T, tanto más cuanto menor sea su aversión a la inflación, b:

$$\frac{dL}{d\pi_t} = 0 \ \Rightarrow\ -2\gamma(\bar{u} - \gamma(\pi_t - \pi_t^e) - u^T) + 2b(\pi_t - \pi^T) = 0$$
$$\xrightarrow{\pi_t^e = \pi_t}\ \pi_1 = \pi^T + \frac{\gamma}{b}(\bar{u} - u^T) > \pi^T$$

CSO: la función de pérdida es una suma de términos cuadráticos convexos y la restricción (curva de Phillips) es lineal en π_t, por lo que L_t es estrictamente convexa en π_t y la CPO es suficiente para un mínimo.

Equilibrio: por tanto, en un entorno de agentes racionales ($\pi_t^e = \pi_t$), el nivel de inflación de equilibrio, π_1, estará por encima del objetivo, π^T. Este exceso, $\frac{\gamma}{b}(\bar{u} - u^T)$, constituye el *sesgo inflacionista* de la política discrecional. Además, dado que $\pi_t^e = \pi_t$, el desempleo se situará en su nivel natural, $\bar{u}$, lo cual anulará cualquier ganancia de bienestar derivada del sesgo inflacionista.

Desarrollo gráfico: el sesgo inflacionista de la política discrecional se representa en el plano tasa de desempleo-inflación (figura A6.2):

1. *Curvas de Phillips a corto plazo*, $\text{CP}_{\text{CP},0}$, $\text{CP}_{\text{CP},0''}$ y $\text{CP}_{\text{CP},1}$: son la restricción del Gobierno, una para cada inflación esperada, con pendiente $-\kappa_u$.

2. *Curva de Phillips a largo plazo*, CP_{LP}: vertical en la tasa natural de paro, $\bar{u}$, donde $\pi_t = \pi_t^e$.

[25] La terminología procede de Backus y Driffill (1985). En su modelo de reputación, el público no sabe si el Gobierno es *wet*, que minimiza la función de pérdida, o *hard-nosed*, comprometido irrevocablemente con una inflación nula, de modo que incluso un Gobierno *wet* puede renunciar a inflar para ganarse esa reputación.

3. *Curvas de indiferencia social*, $\text{CIS}_{0'}$ y CIS_1: son curvas de nivel de la función de pérdida, en torno al punto de máximo bienestar (u^T, π^T). La pérdida es mayor cuanto más alejadas están de él.

4. *Equilibrio*: el punto 1 ($\bar{u}$, π_1) es un equilibrio de Nash, puesto que en él $\text{CP}_{\text{CP},1}$ es tangente a CIS_1, de modo que el Gobierno no tiene incentivos a desviarse, y la inflación coincide con la esperada ($\pi_t = \pi_t^e$). El punto 1 resulta peor que el inicial, porque le corresponde una curva de indiferencia social más distante del punto de máximo bienestar (u^T, π^T). El desempleo es el mismo que al principio, pero la inflación es más alta.

Estática comparativa: el sesgo inflacionista, $\pi_1 - \pi^T = \frac{\gamma}{b}(\bar{u} - u^T)$, depende de tres parámetros del modelo, que solo alteran la inflación (el desempleo de equilibrio sigue siendo $\bar{u}$):

- *Aversión a la inflación*, b: el sesgo disminuye con b y tiende a cero si $b \to \infty$. En ello se funda la delegación en una autoridad más aversa a la inflación que la sociedad (Extensiones).

- *Ambición del objetivo de empleo*, $\bar{u} - u^T$: el sesgo crece con la distancia entre la tasa natural y el objetivo del Gobierno y desaparece si $u^T = \bar{u}$.

- *Pendiente de la curva de Phillips*, $\gamma = 1/\kappa_u$: cuanto más plana es la curva en el plano (u, π) (menor κ_u), más desempleo reduce cada punto de inflación sorpresa, mayor es la tentación de generarla y mayor el sesgo.

Dinámica: la secuencia hipotética que lleva al punto 1 es la siguiente:

1. *Compromiso*: la economía parte del punto 0 ($\pi = \pi^T$, $u = \bar{u}$). La autoridad monetaria anuncia su compromiso de mantener la inflación en su objetivo. Sobre la base de este compromiso, los agentes económicos fijan sus expectativas y firman sus contratos de precios y salarios nominales.

2. *Incentivo a incumplir*: en esta coyuntura, a la autoridad monetaria le conviene romper su promesa a corto plazo, porque, para esos salarios nominales fijos, una inflación mayor que la prevista reduce el salario real, lo cual reducirá el desempleo. Si la autoridad monetaria rompiera su promesa, la economía se situaría en el punto $0'$ ($u_{0'}$, $\pi_{0'}$), donde $\text{CP}_{\text{CP},0}$ es tangente a $\text{CIS}_{0'}$.

3. *Revisión de expectativas*: no obstante, la sociedad incorpora las nuevas expectativas de inflación y exige un salario mayor. Esto provoca un desplazamiento de la curva de Phillips a corto plazo hacia la derecha, a $\text{CP}_{\text{CP},0''}$, hasta el punto $0''$.

4. *Equilibrio de Nash*: el proceso se repetiría hasta llegar al punto 1. La secuencia por etapas es solo expositiva, ya que en el juego único los agentes anticipan el incentivo del Gobierno y la economía se sitúa directamente en ese punto.

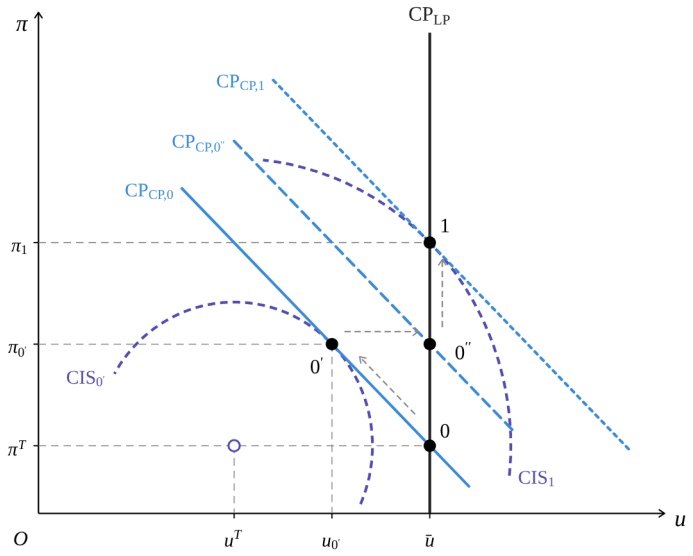

Figura A6.2: Modelo de Barro y Gordon (1983a). Desde el compromiso anunciado (0), el incentivo a incumplir ($0'$) y la revisión de expectativas ($0''$) llevan al equilibrio discrecional (1), con sesgo inflacionista y sin ganancia de empleo.

Extensiones: para la NMC, por tanto, las autoridades monetarias deberían limitarse a controlar la inflación, y se proponen cuatro soluciones al problema de la credibilidad:

1. *Reglas* (Kydland y Prescott, 1977): el banco central queda sujeto a reglas vinculantes (p. ej. un ritmo fijo de expansión monetaria o una tasa de inflación invariable). La regla de Taylor (1993)[26] es una de ellas: $i_t = \bar{i}_t + \alpha_\pi \tilde{\pi}_t + \alpha_y \tilde{y}_t$, con $\bar{i}_t = \bar{r}_t + \pi^T$ y $\tilde{\pi}_t = \pi_t - \pi^T$.

2. *Prácticas reputacionales* (Barro y Gordon, 1983b): si se considera que el juego se repite indefinidamente, la creación de reputación teniendo en cuenta estrategias de castigo puede resolver el problema de la inconsistencia dinámica. Los agentes privados harán demandas de salarios contenidas (bajas expectativas de inflación) si y solo si las autoridades mantienen una inflación baja. Las autoridades monetarias, sabedoras de este comportamiento, pueden estar dispuestas a sacrificar el corto plazo para forjarse una reputación de halcón (actitud dura contra la inflación)[27] y obtener así ganancias en el futuro. En el marco de una economía abierta, esto se puede conseguir a través de, p. ej., un régimen de cambio fijo o una caja de conversión (*currency board*).

3. *Delegación* (Rogoff, 1985): Rogoff defiende delegar la política monetaria en una autoridad independiente con mayor credibilidad. El agente nombrado puede dar un mayor peso al objetivo de estabilidad de precios que el asignado por la sociedad. La mayoría de los países de la OCDE ha seguido esta vía. Desde finales de la década de los ochenta, la evidencia empírica ha mostrado, en las economías avanzadas, una correlación negativa entre inflación e independencia de los bancos

[26] En su calibración original para EE. UU., Taylor (1993) proponía $i_t = \bar{r}_t + \pi_t + 0{,}5\,\tilde{y}_t + 0{,}5\,(\pi_t - \pi^T)$, con $\bar{r}_t = 2\,\%$ y $\pi^T = 2\,\%$, es decir, $\alpha_\pi = 1{,}5$ y $\alpha_y = 0{,}5$. Aunque estas reglas describen razonablemente la conducta de los bancos centrales, la política monetaria óptima rara vez adopta esa forma, pues depende de un estado más complejo de las variables.

[27] Uno de los ejemplos que suele citarse en este contexto es la actuación de la Fed entre 1979 y 1987, durante el mandato de Paul Volcker, que mostró un compromiso ineludible con la estabilidad de precios.

centrales, medida por el grado de atribución de responsabilidades a la entidad independiente. En las economías en desarrollo, en cambio, la independencia legal no guarda esa relación[28].

No obstante, Posen (1993) apunta que correlación no significa causalidad, ya que la independencia y la baja inflación pueden responder a una misma causa, la oposición a la inflación de los grupos con peso político, sobre todo del sector financiero[29]. Además, la delegación introduce costes adicionales ante la existencia de perturbaciones económicas, pues la respuesta estabilizadora de las autoridades puede no ser óptima de acuerdo con las preferencias sociales.

4. *Contratos* (Walsh, 1995): un contrato entre el Gobierno y el banco central puede eliminar el sesgo inflacionista sin renunciar a la respuesta estabilizadora ante las perturbaciones (lo que equivale a replicar la política óptima con compromiso). Para ello, el contrato modifica la función de pérdida de las autoridades monetarias, añadiendo un término de coste adicional, determinado por el contrato entre las dos partes. La cuantificación óptima del coste asociado al contrato vendrá determinada por el sesgo inflacionista que existía en el caso de las políticas discrecionales. Aportaciones posteriores (Svensson, 1997) demuestran que el contrato óptimo de inflación puede implementarse exigiendo a la autoridad monetaria un objetivo de inflación inferior a la tasa socialmente óptima.

Conclusiones:

- *Aspectos positivos*: explica el sesgo inflacionista como resultado de la falta de compromiso, sin suponer agentes ni autoridades irracionales. Asimismo, ofrece el fundamento teórico de las reglas, la independencia del banco central y los objetivos de inflación, y está en la base de la literatura sobre la credibilidad.

- *Limitaciones*: el sesgo descansa en que la autoridad persigue un desempleo inferior al natural, supuesto discutible en un banco central independiente con mandato de estabilidad de precios. El marco es, además, estático y de información completa.

3.2. Implicaciones de política fiscal

En materia de política fiscal, la NMC revitaliza la equivalencia ricardiana, propone la suavización impositiva y advierte de los riesgos de la dominancia fiscal.

3.2.1. La irrelevancia de la financiación del gasto público: revitalización de la equivalencia ricardiana (Barro, 1974)

Origen: Robert Barro, en *Are Government Bonds Net Wealth?* (1974), recupera una tesis que David Ricardo había planteado en el s. XIX, la equivalencia ricardiana[30].

Idea: la aportación de Barro permite ilustrar que los bonos no constituyen riqueza neta de los individuos y, por tanto, la elección entre impuestos y deuda para financiar un gasto público dado es irrelevante.

[28] Suele criticarse también la propia construcción de los indicadores de independencia, que conceden cierto peso a la presencia de la contención de la inflación como objetivo constitutivo de la institución monetaria. Ello conlleva un sesgo en la relación entre las tasas bajas de inflación y la independencia de los bancos centrales.

[29] Una explicación complementaria apela a las preferencias sociales. En Alemania, por ejemplo, la hiperinflación que siguió a la Primera Guerra Mundial generó entre la ciudadanía una aversión muy elevada a la inflación.

[30] Ricardo estudió en su *Essay on the Funding System* (1820) si daba igual financiar una guerra con impuestos o con deuda perpetua, pero dudó de que la equivalencia tuviera consecuencias prácticas, porque creía que los contribuyentes no razonan así. El nombre se generalizó a partir del comentario de Buchanan (1976) al artículo de Barro.

Supuestos:

▪ *Entorno*: mercados de capitales completos y perfectos. Esto supone que el tipo de interés es idéntico para todos los agentes de la economía (p. ej. sector público y agentes privados) y que no existen restricciones de crédito o liquidez.

▪ *Agentes*: intervienen dos:

- *Sector privado*: forma sus expectativas con base en la HER y presenta altruismo intergeneracional perfecto.
- *Sector público*: figuras impositivas neutrales que no distorsionan las decisiones de consumo y ocio (p. ej. impuestos de suma fija), cuyo cambio de calendario no redistribuye renta entre agentes. La senda de gasto público está dada.

Razonamiento: en el modelo walrasiano del ciclo real, al combinar la RPI de los consumidores-productores con la del Gobierno, los bonos desaparecen de la primera[31]. Por ello, para decidir cuánto consumen e invierten, a los hogares solo les importa el valor actual del gasto, que en su cálculo equivale a lo que valen hoy los impuestos.

Los individuos no padecen ilusión de deuda y saben que, tarde o temprano, el Gobierno tendrá que financiar el déficit público de alguna forma. Por tanto, una rebaja de impuestos financiada con deuda (dado el gasto) solo cambiará la composición del ahorro, porque el desahorro público se traducirá en ahorro privado.

La política fiscal no deja por ello de tener efectos reales. Sigue influyendo en el ahorro y, a través de él, en el consumo, en el empleo y en el crecimiento, pero esos efectos no cambian con el modo de financiar el gasto, lo único que la equivalencia ricardiana declara irrelevante.

Extensiones: la consolidación fiscal (política fiscal contractiva) puede tener efectos expansivos o efectos fiscales no keynesianos (Giavazzi y Pagano, 1990), dado que la política fiscal sí tiene efectos reales.

1. *Efecto atracción* (*crowding in*): si los agentes creen que el recorte del gasto público es permanente, anticipan una carga tributaria futura más baja. Su riqueza en valor presente crece y su incertidumbre disminuye, lo que estimula el consumo y la inversión.

2. *Primas de riesgo*: la consolidación también reduce las de la deuda pública.

La equivalencia ricardiana sigue valiendo en una economía abierta. Si se bajan los impuestos y el gasto público se mantiene, la deuda que emite el Estado la absorbe el ahorro adicional de los hogares, sin que varíe el saldo por cuenta corriente.

Conclusiones:

▪ *Aspectos positivos*: la hipótesis de equivalencia ricardiana entre deuda e impuestos es un valioso punto de partida teórico.

▪ *Limitaciones*: sin embargo, hay pocos motivos para esperar que sea una buena primera aproximación al comportamiento real de la economía, dado que sus supuestos rara vez se cumplen:

[31] Esquema. (i) Ecuación de estado de los hogares con sector público: $c_t + (k_{t+1} - k_t) + (b_{t+1} - b_t) = y_t - T_t + r_t b_t$. (ii) Aplicando recursividad hacia delante y la condición de no Ponzi se obtiene la restricción presupuestaria intertemporal de los hogares en valor presente. (iii) Haciendo lo propio con la restricción del sector público, $(1+r)b_t = \sum_{j\geq 0}(1+r)^{-j}(T_{t+j} - G_{t+j})$. (iv) Sustituyendo esta en aquella, los bonos públicos b_t desaparecen de la restricción de los hogares, por lo que no son riqueza neta.

- *Mercados de capitales*: generalmente existen restricciones de liquidez e imperfecciones. Por ejemplo, el Gobierno ofrece más garantías que los prestatarios privados y por eso suele endeudarse a un tipo de interés más bajo. En esas condiciones, la deuda pública sí es riqueza.

- *Agentes*: por un lado, su racionalidad es limitada. La teoría de la ilusión fiscal de Buchanan y Wagner (1977) sostiene que los agentes pueden subestimar los impuestos que pagan. El sector público puede entonces crecer sin que noten que la presión fiscal sube en la misma medida. Por otro, cada individuo vive menos que la economía y su altruismo hacia las generaciones siguientes puede faltar o ser incompleto, de modo que la carga impositiva puede trasladarse a generaciones futuras.

- *Estado*: los impuestos rara vez son de suma fija, sino distorsionantes (afectan a las decisiones intertemporales).

3.2.2. Suavización impositiva (Barro, 1979)

Origen: si los impuestos son distorsionantes, la forma de financiación del gasto deja de ser irrelevante y queda por determinar cuál es la senda óptima de impuestos y deuda. Robert Barro (1979) argumenta que la determinación de la política fiscal óptima se puede plantear como un programa de optimización intertemporal.

Idea: lo óptimo es mantener constante el tipo impositivo esperado (*tax smoothing*), recomendación central de la aportación de Barro en materia de política fiscal. Con incertidumbre, el tipo sigue un paseo aleatorio y solo cambia con la información nueva sobre el gasto[32].

Razonamiento: en ese programa, dada la senda de gasto público, el sector público elige la senda de impuestos (y, con ella, de deuda) que hace mínima la pérdida de bienestar y cumple su restricción presupuestaria intertemporal, así como las condiciones de optimalidad del sector privado.

Dado que el coste marginal de la distorsión aumenta con el tipo impositivo, mantener un tipo uniforme reduce al mínimo el exceso de gravamen. De ello se sigue que es preferible:

1. *Impuestos bajos*: mantenerlos durante muchos años antes que impuestos altos durante uno solo.

2. *Estabilizadores automáticos*: dejar fijo el tipo impositivo y que los estabilizadores automáticos generen déficit en las recesiones y superávit en las expansiones. El déficit resulta así contracíclico y el ahorro público, procíclico. Al absorber los *shocks*, la deuda permite suavizar los impuestos a lo largo del tiempo.

Conclusiones:

▪ *Aspectos positivos*: ofrece una teoría de la deuda pública óptima, según la cual los déficits y superávits cíclicos pueden responder a una política óptima y no a falta de disciplina. Justifica así que actúen los estabilizadores automáticos y que los aumentos transitorios del gasto se financien con deuda.

▪ *Limitaciones*: supone un Gobierno benevolente, con horizonte largo y una senda de gasto dada, por lo que no explica la acumulación persistente de deuda observada en muchos países, que responde a factores políticos ajenos al modelo.

[32] Lucas y Stokey (1983) muestran que, si el Gobierno puede emitir deuda contingente al estado de la economía, el tipo óptimo hereda la autocorrelación del gasto y no sigue un paseo aleatorio, aunque sigue siendo más estable que con un presupuesto equilibrado en cada periodo.

3.2.3. Dominancia fiscal: la aritmética monetarista desagradable (Sargent y Wallace, 1981)

Origen: un último resultado, en la frontera entre la política fiscal y la monetaria, se debe a Sargent y Wallace, en *Some Unpleasant Monetarist Arithmetic* (1981).

Idea: se habla de *aritmética monetarista desagradable* cuando la política monetaria resulta incapaz de controlar la inflación a causa de una política fiscal sistemáticamente expansiva, financiada mediante deuda, que acaba siendo más inflacionista que su monetización.

Razonamiento: estos autores parten de la restricción presupuestaria intertemporal del sector público, según la cual un déficit primario persistente ha de financiarse con deuda o con creación de dinero (señoraje). Si la autoridad monetaria decide ser restrictiva hoy (no monetizar), el déficit se financia íntegramente con deuda, lo que pone en marcha un efecto bola de nieve:

1. $r > g_Y$: si el tipo de interés real de la deuda supera la tasa de crecimiento de la economía ($r > g_Y$), la ratio deuda/PIB crece de forma explosiva, pues los intereses han de financiarse con nuevas emisiones.

2. *Límite de endeudamiento*: llegado un punto, el mercado se niega a seguir absorbiendo deuda pública, dados los riesgos de impago, o exige primas crecientes que aceleran el proceso.

3. *Dominancia fiscal*: alcanzado ese límite, y si la política fiscal no se corrige (dominancia fiscal), la autoridad monetaria no tiene más remedio que monetizar el déficit y el servicio de la deuda acumulada. La inflación futura será, así, mayor que la que habría resultado de monetizar desde el principio.

Bajo la HER, los agentes racionales anticiparán la monetización futura. En consecuencia, la inflación se acelerará ya hoy (p. ej. vía reivindicaciones salariales y demanda de dinero), de modo que una política monetaria restrictiva puede resultar, paradójicamente, más inflacionista que una acomodaticia, y de ahí lo desagradable de la aritmética.

Por ello, existe un círculo vicioso entre la política fiscal sistemáticamente expansiva financiada vía deuda y la inflación, con lo que esta pasa a tener un origen más fiscal que monetario.

Conclusiones:

- *Aspectos positivos*: muestra que la política monetaria y la fiscal están ligadas por la restricción presupuestaria intertemporal del sector público, de modo que el control de la inflación exige disciplina fiscal. Asimismo, la teoría fiscal del nivel de precios generaliza su razonamiento (tema A39, apartado 3.2.2), y el modelo ofrece un argumento a favor de las reglas fiscales.

- *Limitaciones*: el resultado depende de supuestos concretos, como que el tipo de interés real supere al crecimiento, que exista un límite a la deuda que el público está dispuesto a absorber y, sobre todo, que la autoridad fiscal no ajuste. Además, el modelo supone, en lugar de explicar, cuál de las dos autoridades cede.

3.3. Críticas al pensamiento de la NMC

Las críticas a la NMC se dirigen, sobre todo, a la HER y al vaciado continuo de los mercados.

3.3.1. Críticas a la HER

A pesar de su capacidad explicativa y rigor conceptual, la hipótesis de expectativas racionales no ha estado exenta de críticas. Destacan tres:

- *Asimetrías de información*: la formulación canónica de la HER pasa por alto que unos agentes están mejor informados que otros.

- *Costes de información*: la HER supone gratuita la información, pues solo así compensaría a los agentes aprovecharla toda, y el supuesto es poco realista[33].

- *Riesgo frente a incertidumbre*: la HER atribuye a los agentes el conocimiento de la distribución de probabilidad de cada variable relevante, sobre la que calculan una esperanza condicionada. El supuesto corresponde a una situación de *riesgo* y excluye la *incertidumbre* en sentido de Knight. La diferencia importa porque, según muchos autores, lo que dominó la crisis financiera global (2008-2014) fue la incertidumbre y no el riesgo. Por eso la NMC no ha logrado explicar la crisis de forma plenamente satisfactoria[34].

3.3.2. Crítica al vaciado continuo de los mercados

El pensamiento de la NMC descansa en el vaciado continuo gracias a la plena flexibilidad de precios (apartado 1.1), pero en la realidad existen rigideces de precios y salarios. Por ello, escuelas de pensamiento posteriores han incorporado este tipo de fricciones.

En concreto, a comienzos de los años ochenta un nuevo grupo de economistas[35] dio forma a la **nueva economía keynesiana**, que introdujo en el estudio del mercado laboral y del de bienes la competencia imperfecta, junto con fricciones y rigideces tanto nominales como reales[36].

3.4. La nueva síntesis neoclásica

Las escuelas acabaron convergiendo en un nuevo consenso cuyos principios se recogen después.

3.4.1. La convergencia del pensamiento económico

Según el economista estadounidense Michael Woodford (2009), desde finales de la década de 1990 comienza a observarse una mayor convergencia entre las opiniones de los diferentes macroeconomistas.

Si bien todavía existen problemas que están lejos de ser resueltos, ya no existen las diferencias tan fundamentales que existían en épocas anteriores del pensamiento económico (p. ej. en cuanto

[33] Como alternativa, Feige y Pearce (1976) proponen las expectativas *económicamente racionales*, que tienen en cuenta el coste de la información, de forma que los agentes hacen uso de la información mientras el Bmg sea mayor que el Cmg.

[34] Las respuestas posteriores mantienen la optimización, pero relajan la información que se atribuye a los agentes. En el aprendizaje adaptativo, los agentes estiman el modelo con los datos y pueden converger a las expectativas racionales (Evans y Honkapohja, 2001), y en la inatención racional procesar la información tiene un coste (Sims, 2003).

[35] Entre los más destacados se encuentran George Akerlof, Ben Bernanke, Janet Yellen, Olivier Blanchard, Gregory Mankiw, Lawrence Summers o Stanley Fischer.

[36] La NEK aceptó, con matices, varios supuestos de la NMC. Sus agentes tienen expectativas racionales y no sufren ilusión monetaria, hay una tasa natural de paro y el marco es de equilibrio general dinámico, estocástico y microfundamentado. Rechazó, en cambio, la competencia perfecta y el vaciado continuo de los mercados, y con frecuencia las hipótesis sobre la información.

a la metodología[37]). Además, la línea que tradicionalmente separaba la microeconomía y la macroeconomía es cada vez más difusa.

El principal motivo de la convergencia es que ha sido necesario realizar concesiones de un lado y de otro en vista de la mera observación empírica y la experiencia, así como la mayor profundidad de las propias ideas que forman parte del pensamiento económico.

3.4.2. Los principios del nuevo consenso (Woodford, 2009)

Siguiendo a Marvin Goodfriend y Robert King (1997), se puede hablar de una **nueva síntesis neoclásica** o, según otros autores, de un "nuevo consenso". El nuevo consenso descansa en varios principios: tres relativos a la metodología, uno al impulso del ciclo y uno en materia de política económica.

Metodología:

- *Equilibrio general estocástico*: existe consenso en que el análisis macroeconómico debe apoyarse en modelos de equilibrio general estocástico intertemporal, que permiten entender tanto las fluctuaciones a corto plazo como el crecimiento a largo plazo en un marco integrado. Este tipo de modelos ha fusionado las aportaciones de la nueva macroeconomía clásica (uso de modelos del ciclo real) con las de la nueva economía keynesiana (rigideces en los mercados de bienes y de trabajo).

- *Validez estructural*: existe consenso en cuanto a la conveniencia de basar los análisis cuantitativos de política económica en modelos econométricos robustos a la crítica de Lucas.

- *Expectativas endógenas*: existe consenso en que las expectativas deben determinarse dentro del propio modelo. En concreto, los modelos adoptan la HER como teoría razonable de formación de expectativas, sin que ello lleve necesariamente a las implicaciones de política que inicialmente se derivaron de ella (la inefectividad de la política económica anticipada).

Impulso del ciclo: la nueva síntesis acepta que los *shocks* reales (p. ej. tecnología, oferta, preferencias o *shocks* fiscales) son una fuente importante de fluctuaciones económicas. En cambio, se estima que los *shocks* de política monetaria explican solo una pequeña parte de la varianza de las fluctuaciones del PIB.

Política económica:

- *Política monetaria*: se considera efectiva, en general, y en especial como medio para controlar la inflación, dado que tiene efectos reales. Se abandona así la dicotomía clásica y la neutralidad monetaria a corto plazo.

- *Política fiscal*: no figura entre esos principios, pero tras la Gran Recesión se le reconoce un papel como estabilizador, sobre todo cuando los tipos de interés están en el ZLB.

Gráficamente, en los planos brecha de producción-desviación del tipo de interés real respecto del natural y brecha de producción-inflación (figuras A6.3 y A6.4), la curva IS refleja la demanda agregada. Tiene pendiente negativa porque un tipo de interés real superior al natural reduce el gasto y abre una brecha de producción negativa. La curva de Phillips, CP, recoge la inflación que genera cada brecha y es creciente, porque más actividad eleva los costes y los precios. La curva de regla monetaria (MR) resume la respuesta del banco central y es decreciente, porque, ante más inflación, el banco central sube el tipo de interés real y reduce la producción.

[37] Por supuesto, se refiere al pensamiento ortodoxo, en la medida en que muchas escuelas heterodoxas (p. ej. austriacos) no comparten el formalismo matemático que existe hoy en día en la economía.

La economía parte del punto 0, con la inflación en el objetivo, π^T, y brecha de producción nula. Ante un *shock* tecnológico positivo ($\Delta a > 0$), si el tipo de interés no sigue la caída del natural, la MR se desplaza de MR_0 a MR_1 y la economía pasa del punto 0 al 1. Ante un *shock* de costes ($\varepsilon^\pi > 0$), la curva de Phillips se desplaza de CP_0 a CP_1 y el banco central acepta una brecha de producción negativa (del punto 0 al 1) para devolver la inflación al objetivo.

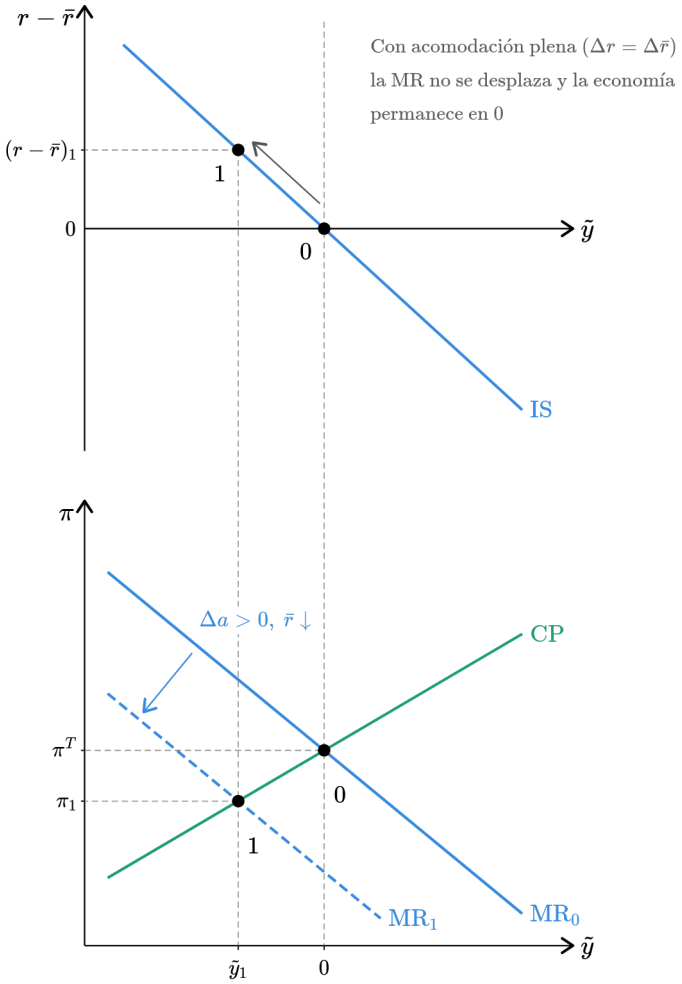

Figura A6.3: Modelo IS-PC-MR ante un *shock* tecnológico positivo ($\Delta a > 0$). Si la regla de Taylor no sigue la caída del tipo de interés natural, el tipo real queda por encima del natural, la MR se desplaza a la izquierda y la economía pasa de 0 a 1, con brechas de producción e inflación negativas. Con acomodación plena, la brecha no se abre y la economía permanece en 0.

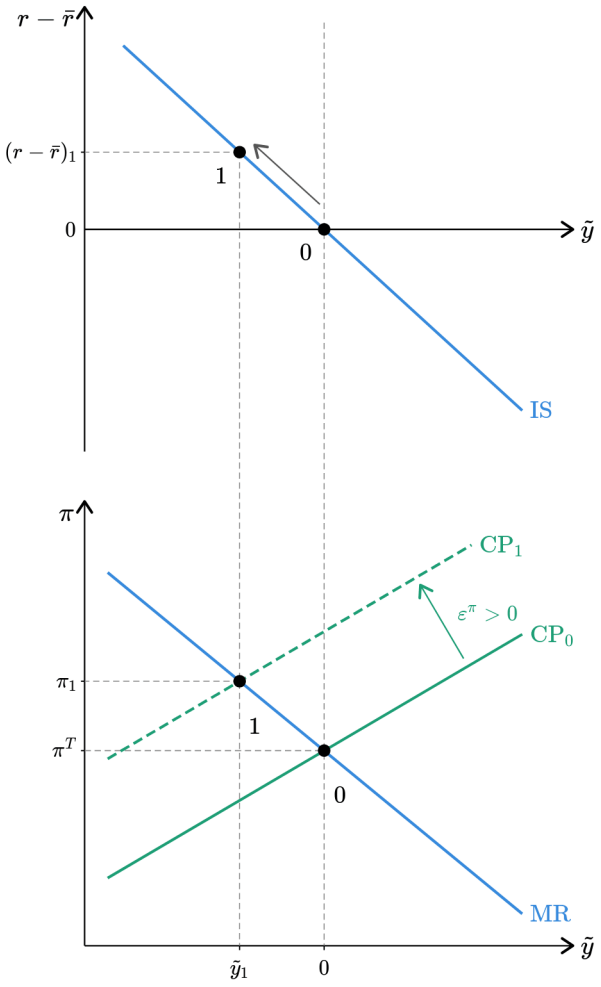

Figura A6.4: Modelo IS-PC-MR ante un *shock* de costes. El banco central acepta transitoriamente una brecha de producción negativa para devolver la inflación al objetivo.

El modelo canónico de la nueva síntesis (IS-PC-MR) y su análisis gráfico se desarrollan en el tema A07, apartado 2.1, y las extensiones que responden a la carencia de sector financiero de estos modelos (apartado 2.3 de este tema), en el apartado 3.4.2 de ese tema.

Conclusiones

La nueva macroeconomía clásica cambió el método de la disciplina. La crítica de Lucas obligó a derivar las relaciones agregadas de agentes que optimizan con expectativas racionales. Con ese método y el supuesto de mercados que se vacían, la escuela explicó el ciclo como respuesta óptima a sorpresas monetarias (Lucas) o a perturbaciones tecnológicas (ciclo real). Extrajo, además, conclusiones de política radicales, como la inefectividad de la política monetaria anticipada, el sesgo inflacionista de la discrecional, la irrelevancia de la financiación del gasto público y la suavización de los impuestos. Sus críticos cuestionaron el vaciado continuo de los mercados y la HER, pero aceptaron su método, y el debate posterior se libró con sus reglas.

El tema se sitúa entre dos escuelas. Parte de la curva de Phillips ampliada por las expectativas, en la que la NMC sustituye las expectativas adaptativas por las racionales (tema A05), y desemboca en la nueva economía keynesiana, que introdujo rigideces y competencia imperfecta en el marco de la NMC y, con la nueva síntesis, devolvió a la política monetaria sus efectos reales a corto plazo (tema A07). Sus resultados tienen, además, largo recorrido. La teoría del ciclo real es una de las explicaciones de equilibrio de las fluctuaciones (tema A42). La inconsistencia dinámica fundamenta el debate entre reglas y discrecionalidad y la independencia de los bancos centrales (tema A36), y la aritmética monetarista desagradable muestra que la política monetaria no controla la inflación si la fiscal no garantiza la solvencia (tema A39).

La Gran Recesión abrió el debate sobre los modelos DSGE. Sus defensores los consideran la herramienta principal para evaluar políticas de forma transparente, ampliable con fricciones financieras y agentes heterogéneos (Christiano, Eichenbaum y Trabandt, 2018). Los críticos replican que la microfundamentación de estos modelos omitía la información imperfecta, el sistema financiero y la distribución de la renta (Stiglitz, 2018). Entre ambas posturas, Blanchard (2018) defiende la convivencia de varias clases de modelos (tema A01, apartado 3.4.3). Sigue abierta, además, la discusión sobre la propia HER. Las encuestas de previsiones muestran errores predecibles, compatibles con rigideces de información en las previsiones medias (Coibion y Gorodnichenko, 2015) y con sobrerreacción en las individuales (Bordalo, Gennaioli, Ma y Shleifer, 2020). El repunte inflacionista posterior a la pandemia ha llevado la cuestión a la política monetaria, porque, si las expectativas se desanclan, solo una recesión profunda devuelve la inflación al objetivo (Reis, 2021; tema A37, apartado 2.4.2).

Anexos

A.1. La parábola de las islas de Phelps y la curva de oferta de Lucas

La curva de oferta de Lucas del apartado 2.1 entronca con la parábola de las islas de Phelps, metáfora de la información incompleta.

Los agentes se distribuyen aleatoriamente en islas (mercados). Cada oferente produce un único bien, consume muchos y decide su producción con base en los precios relativos. Cada productor conoce solo el precio de su producto, p_{it}, pero no el nivel general de precios, p_t, por lo que tiende a confundir lo que ocurre en su mercado con lo que ocurre en la economía. Su oferta es creciente en el precio relativo (todo en logaritmos):

$$y_{it} = \bar{y} + \eta\,(p_{it} - p_t), \qquad \eta > 0$$

Como no conoce p_t, lo estima según la HER mediante $\mathbb{E}[p_t \mid p_{it}]$. Al observar una subida de p_{it}, no sabe si responde a un *shock* específico (de su mercado) o general (agregado), por lo que realiza una extracción de señal asignando una probabilidad ϑ a que sea específico (apartado 2.1.2). Con los *shocks* distribuidos como

$$\text{específico: } z_{it} \sim N(0, \sigma_i^2), \qquad \text{general: } p_t \sim N(\bar{p}, \sigma_p^2), \qquad p_{it} = p_t + z_{it},$$

su expectativa sobre el nivel general es una media ponderada del *a priori* y de la señal que aporta su propio precio:

$$\mathbb{E}[p_t \mid p_{it}] = \vartheta\,\mathbb{E}[p_t] + (1 - \vartheta)\,p_{it}, \qquad \vartheta = \frac{\sigma_i^2}{\sigma_i^2 + \sigma_p^2}$$

Sustituyendo la estimación en la oferta individual, esta responde al precio relativo percibido:

$$y_{it} = \bar{y} + \eta\,(p_{it} - \mathbb{E}[p_t \mid p_{it}]) = \bar{y} + \vartheta\eta\,(p_{it} - \bar{p}) \equiv \bar{y} + \zeta\,(p_{it} - \bar{p})$$

Agregando islas idénticas (el componente específico z_{it} promedia a cero), con $p_t^e \equiv \mathbb{E}[p_t] = \bar{p}$, se recupera la oferta agregada de Lucas del cuerpo del tema:

$$y_t = \bar{y} + \zeta\,(p_t - p_t^e)$$

A.2. La curva de Phillips con HER

A partir de la oferta agregada de Lucas puede obtenerse una curva de Phillips ampliada por las expectativas, aquí racionales (su tratamiento en el marco de la teoría del desempleo se encuentra en el tema A28). Reordenando la OA, la brecha de producción depende de la sorpresa de precios. Después, expresando esta en tasas ($\pi_t = p_t - p_{t-1}$) y traduciendo la brecha de producción a brecha de desempleo mediante la ley de Okun (1962), $\tilde{y}_t = -\phi\,(u_t - \bar{u})$, con $\phi > 0$ el coeficiente de Okun, se llega a:

$$\pi_t = \pi_t^e - \kappa_u\,(u_t - \bar{u}) - \frac{\varepsilon_t^s}{\zeta}, \qquad \kappa_u = \phi/\zeta > 0$$

Es decir, la inflación depende de la inflación esperada, de la desviación del desempleo respecto de su tasa natural y de las perturbaciones de oferta. En el largo plazo, cuando $\pi_t = \pi_t^e$, el desempleo se sitúa en su tasa natural $\bar{u}$ (curva de Phillips vertical en el plano (u, π)).

A7. La nueva economía keynesiana

La nueva economía keynesiana. Primera y segunda generación.

Introducción

> **Introducción común.** Los temas A02 a A07 comparten una introducción, que se desarrolla al comienzo del tema **A02**. Recorre la evolución del pensamiento económico, del mercantilismo y la fisiocracia a la escuela clásica, la revolución marginalista, Keynes y la síntesis neoclásica, el monetarismo, la nueva macroeconomía clásica, la nueva economía keynesiana y la nueva síntesis neoclásica.

La nueva macroeconomía clásica (tema A06) dejó a los keynesianos una pregunta difícil de eludir, la de por qué no se ajustan los precios y los salarios hasta vaciar los mercados si los agentes son racionales y optimizan. La **nueva economía keynesiana** (NEK) nació en los años ochenta para responderla sin renunciar al método de sus rivales.

Sus autores, entre ellos Mankiw, Akerlof, Yellen, Blanchard, Stiglitz, Fischer, Taylor y Calvo, aceptan las expectativas racionales y la optimización, pero introducen **competencia monopolística**, información imperfecta y **rigideces nominales y reales** que explican por qué los precios se ajustan con lentitud. Los contratos escalonados de Fischer (1977) y Taylor (1980) mostraron que, aun con expectativas racionales, la política monetaria puede estabilizar la economía. Por su parte, los costes de menú de Mankiw (1985) y la cuasirracionalidad de Akerlof y Yellen (1985) probaron que fricciones privadas pequeñas pueden tener grandes efectos agregados, que las rigideces reales amplifican. Con precios rígidos, el ajuste se produce por cantidades, la política monetaria tiene efectos reales y la economía puede quedar atrapada en **fallos de coordinación**. De ahí que la NEK justifique la intervención para corregir los fallos de mercado y una política de estabilización basada en reglas.

La primera generación injertó esas fricciones en el análisis de equilibrio parcial heredado del modelo IS-LM. La segunda, desde los años noventa, las integró en modelos de equilibrio general dinámico con la fijación escalonada de precios de Calvo (1983). El resultado es el **modelo de tres ecuaciones**, con una curva IS dinámica, una curva de Phillips neokeynesiana y una regla de política monetaria como la de Taylor (1993), que con la nueva síntesis neoclásica (Woodford, 2003; Galí, 2008) es el marco de referencia de los bancos centrales. El repunte inflacionista de 2021-2023 ha puesto a prueba ese marco y ha reavivado el debate sobre el papel de las expectativas y de las perturbaciones de oferta.

Las aportaciones de la NEK se exponen en este orden:

- *Fundamentos*: los rasgos de la NEK y las teorías de la primera generación sobre las rigideces reales y nominales, cuyo análisis dinámico desemboca en la fijación de precios de Calvo que adopta la segunda generación.

- *Desarrollos posteriores*: el modelo canónico de la segunda generación (IS-PC-MR), que permite estudiar el ciclo, y sus extensiones.

- *Debate*: las implicaciones de política monetaria, fiscal y de empleo del pensamiento de la NEK, sus críticas y la nueva síntesis neoclásica.

1. Fundamentos del pensamiento neokeynesiano

Los autores neokeynesianos son heterogéneos, pero sus modelos coinciden en justificar la posible existencia de equilibrios subóptimos a corto plazo que requieren la intervención de las autoridades.

1.1. Principales rasgos

Metodología: depende de la generación de autores. La primera y la segunda generación adoptan, en general, la hipótesis de expectativas racionales (HER), lo que las separa de la tradición poskeynesiana. La primera generación (Fischer, 1977; Taylor, 1980; Mankiw, 1985) se apoya todavía en el análisis de equilibrio parcial y en herramientas heredadas del paradigma IS-LM, injertando las fricciones sobre dicho marco. Sin embargo, la crítica de Lucas (1976) impulsó el programa hacia la microfundamentación en equilibrio general. La segunda generación completa ese movimiento con la adopción de modelos de equilibrio general dinámico y estocástico (DSGE), en el marco de la nueva síntesis neoclásica (Woodford, 2003; Galí, 2008).

Supuestos:

▪ *Entorno*: la NEK enfatiza el papel de las características institucionales en la conducta de los agentes. Si bien sus autores apoyan la microfundamentación de la macroeconomía, también insisten en la macrofundamentación de la microeconomía.

Las instituciones restringen la conducta individual, porque muchos intercambios nacen de contratos que deben respetarse en el tiempo (p. ej. relaciones de agencia).

Todo esto lleva, a diferencia del pensamiento de corte clásico, a la consideración de fallos de mercado o imperfecciones. Generalmente, estas incluyen:

- *Poder de mercado*: en el mercado de bienes y servicios y de factores. Generalmente se supone competencia monopolística[1], si bien también existen modelos basados en la teoría de oligopolios y de monopolio.

- *Información imperfecta*: especialmente, información asimétrica en diferentes industrias (p. ej. mercados financieros o, en general, las relaciones contractuales).

- *Rigideces de precios*: a diferencia del mundo clásico, en el que existe plena flexibilidad de precios (ya sean de bienes, salarios o tipos de interés), en el mundo neokeynesiano habitualmente existen rigideces, que pueden ser nominales o reales:

    - *Rigideces nominales*: resistencia al ajuste de los precios y salarios nominales. Su fundamento son fricciones nominales (costes de menú, contratos escalonados…) que, pese a implicar pérdidas privadas de segundo orden para la empresa, generan efectos agregados de primer orden.
    - *Rigideces reales*: baja sensibilidad del precio (o salario) *real deseado* (el óptimo de la empresa) ante variaciones de la actividad agregada. Surgen de factores estructurales como la aversión al riesgo vía contratos implícitos, los salarios de eficiencia o la curvatura de la demanda derivada de las preferencias. Su papel es amplificar la rigidez nominal (Ball y Romer, 1990). Como el óptimo apenas se desplaza, la pérdida de no ajustar el precio nominal es pequeña, lo que ensancha el intervalo de inacción.

[1] En un trabajo de 1986, Hall resumió de forma estilizada lo que había observado en cincuenta industrias. Las empresas fijaban precios superiores a su coste marginal (aplicaban, por tanto, un *mark-up*) y producían donde su coste medio todavía decrecía, pero apenas obtenían beneficios económicos, como en el equilibrio de Chamberlin con libre entrada. Esos resultados encajan con un mercado de competencia monopolística.

Una de las implicaciones de este tipo de rigideces es la existencia de *fallos de coordinación*. Como no existe plena flexibilidad de precios y salarios, pueden darse situaciones de equilibrios múltiples subóptimos, en las que un grupo de agentes podría conseguir una mejor solución, pero no lo consigue porque no se coordina[2]. El mecanismo subyacente es la complementariedad estratégica en la fijación de precios (Cooper y John, 1988), la misma que origina las rigideces reales. De ahí pueden surgir "profecías autocumplidas", en las que el optimismo o el pesimismo individual produce, por agregación, el resultado esperado[3].

- *Agentes*: el grueso de la NEK mantiene la racionalidad plena y la formación de expectativas con base en la HER. Una parte de la literatura admite, sin embargo, desviaciones menores de la optimización (la *near-rationality* de Akerlof y Yellen, 1985) o sesgos cognitivos que limitan la racionalidad de los agentes[4].

Ajuste: de todo lo anterior se deriva que no hay un ajuste instantáneo ni un vaciado continuo de los mercados, sino lentitud en el ajuste, que se realiza preferentemente vía cantidades. Con los precios rígidos, la producción y el empleo absorben las perturbaciones, mecanismo heredado de los modelos de desequilibrio o *fix-price* (Clower, 1965; Barro y Grossman, 1971; Malinvaud, 1977).

Política económica: se aboga por la intervención para corregir los fallos de mercado. Además, como las rigideces nominales rompen la neutralidad monetaria, la política monetaria sistemática tiene efectos reales, lo que explica el énfasis posterior en las reglas de tipos de interés y en los objetivos de inflación (principio de Taylor, 1993).

1.2. Teorías sobre rigideces reales

Una línea de investigación de la NEK de primera generación, paralela a la de las rigideces nominales, estudió las rigideces e imperfecciones reales. Estas teorías explican por qué el salario o el precio real deseado apenas responde a la actividad, primero en el mercado de trabajo y después en el de bienes y servicios.

1.2.1. Rigideces reales en el mercado de trabajo

En el mercado de trabajo, cuatro teorías explican por qué el salario real puede diferir del de vaciado.

1.2.1.1. Teoría de los contratos implícitos (Baily, 1974; Gordon, 1974; Azariadis, 1975)

Origen: en un mercado de trabajo walrasiano las empresas contratarían personal nuevo en cada periodo, pero la evidencia revela relaciones laborales duraderas, casi vitalicias para una fracción importante de los trabajadores de EE. UU. (Hall, 1982), algo que ese modelo no contempla.

[2] En teoría de juegos, los juegos de coordinación como la caza del ciervo (*stag hunt*) ilustran este fenómeno, porque tienen equilibrios múltiples ordenados en el sentido de Pareto y los agentes pueden quedar atrapados en el inferior. No así el dilema del prisionero, cuyo equilibrio (ineficiente) es único.

[3] Por ejemplo, si una empresa anticipa una recesión y despide empleados, reduce la demanda de otras empresas y desencadena una reacción en cadena que confirma la recesión.

[4] El concepto de racionalidad limitada (*bounded rationality*) procede de H. A. Simon (1957). Por los costes de recoger y procesar información, los agentes aplican heurísticas (reglas simples e intuitivas) en lugar de reglas estrictas de optimización.

Las teorías de los contratos implícitos se apoyan en este hecho para explicar por qué los salarios pueden diferir de los de vaciado ante determinados fallos de mercado como la información imperfecta.

Idea: dentro de una relación laboral a largo plazo, la renta salarial actúa como un seguro para el trabajador, por lo que no se ajusta a las fluctuaciones de la demanda de trabajo.

Razonamiento: bajo incertidumbre, con trabajadores aversos al riesgo y empresas neutrales al riesgo, los contratos de largo plazo fijan rentas salariales estables que aseguran al trabajador frente a una demanda de trabajo fluctuante. Así, la renta salarial no obedece a las fuerzas de mercado del periodo (se desvincula de la productividad corriente), si bien el empleo puede seguir fijándose de forma eficiente. El reparto óptimo del riesgo fija una renta menor en los estados favorables que el valor de lo producido por el trabajo[5].

Conclusiones:

- *Aspectos positivos*: explica la escasa sensibilidad de la renta salarial real al ciclo por el reparto óptimo del riesgo, sin recurrir a rigideces nominales ni a la irracionalidad de los agentes.

- *Limitaciones*: la teoría explica mejor la estabilidad de la renta salarial que el desempleo involuntario, porque el empleo puede seguir siendo eficiente.

1.2.1.2. Teoría de los salarios de eficiencia (Solow, 1979; Shapiro y Stiglitz, 1984)

Origen: el concepto de *salario de eficiencia* fue introducido ya por Marshall (1890) para reflejar el salario pagado por la eficiencia de cada unidad de trabajo[6]. Las teorías modernas de la NEK explican por qué los salarios pueden superar los de competencia perfecta.

Idea: el salario afecta a la productividad (esfuerzo) del trabajador, por lo que a las empresas les interesa pagar salarios superiores a los de vaciado.

Razonamiento: en el modelo de Shapiro y Stiglitz[7], hay información asimétrica poscontractual en la relación laboral, porque el esfuerzo del trabajador no es observable. Las empresas lo incentivan, por ello, con salarios superiores a los de vaciado.

Conclusiones:

- *Aspectos positivos*: explica el desempleo involuntario como un equilibrio en el que empresas que maximizan beneficios no bajan el salario aunque haya trabajadores dispuestos a trabajar por menos, y aporta una fuente de rigidez del salario real.

- *Limitaciones*: supone que las empresas no disponen de otros mecanismos de incentivo (p. ej. salarios diferidos que crecen con la antigüedad) que harían innecesario pagar un salario superior al de vaciado.

[5] Que la renta salarial quede por debajo de lo producido en los estados favorables es la prima del seguro. El trabajador averso al riesgo lo acepta a cambio de estabilidad, y la empresa neutral al riesgo solo ha de cubrir la utilidad de reserva del trabajador.

[6] Para Marshall, cada trabajador se remunera según su eficiencia, de modo que el salario por unidad de eficiencia se iguala y el empleador es indiferente entre trabajadores más y menos eficientes.

[7] Esta es una de las teorías que explican la existencia de salarios de eficiencia: nutrición, rotación (si no pagan bien, los trabajadores se van), selección (mayores salarios atraen mejores trabajadores), evitar la elusión del esfuerzo (*shirking*, es decir, el riesgo moral) y motivos sociológicos (mejora de la moral).

1.2.1.3. Teoría de los *insiders* y *outsiders* (Lindbeck y Snower, 1988)

Origen: la teoría de los trabajadores internos y externos se debe a Assar Lindbeck y Dennis Snower, que la desarrollan a lo largo de los años ochenta.

Idea: los trabajadores empleados (*insiders*) negocian los salarios en su propio interés, sin tener en cuenta a los desempleados (*outsiders*), por lo que el salario se sitúa por encima del de vaciado.

Razonamiento: estos modelos consideran trabajadores heterogéneos en sus preferencias y en su peso en las negociaciones salariales. Las empresas mantienen así relaciones con dos grupos, los *insiders* (con los que tienen relaciones específicas y cuyos intereses suelen recogerse en convenios) y los *outsiders* (trabajadores no contratados que pueden serlo tras la firma del convenio). Ello genera una asimetría de poder de negociación en favor de los *insiders* (especialmente con sindicatos fuertes)[8].

El resultado es triple:

- *Dualidad* (marginación de los *outsiders*): existe un porcentaje considerable de desempleados (es decir, *outsiders*) sin influencia sobre la fijación de salarios.

- *Rigidez salarial (anticíclica bajo ciertas condiciones)*: en fases recesivas se despide a parte de la plantilla y los *insiders* que permanecen fijan el salario atendiendo a su propio empleo y no al de la economía. De ahí que el salario real apenas caiga pese al aumento del paro e incluso, bajo ciertas condiciones, suba.

- *Histéresis*: perturbaciones transitorias pueden tener efectos permanentes sobre el desempleo (véase el tema A28, apartado 2.3).

Conclusiones:

- *Aspectos positivos*: explica el desempleo involuntario y la rigidez del salario real por el poder de negociación de los trabajadores empleados, sin recurrir a rigideces nominales. Aporta, además, uno de los mecanismos de persistencia del desempleo en que se apoya la hipótesis de histéresis.

- *Limitaciones*: el resultado descansa en que los costes de rotación laboral (despido, contratación y formación) sean elevados, pues solo así las empresas no sustituyen a los *insiders* por *outsiders* dispuestos a trabajar por un salario menor. Su alcance es, por ello, mayor en mercados con sindicatos fuertes y alta protección del empleo.

1.2.1.4. Modelo macroeconómico de negociación en el mercado de trabajo (Layard y Nickell, 1985)

Origen: Richard Layard y Stephen Nickell (1985) presentan su macromodelo de negociación colectiva[9], y Wendy Carlin y David Soskice lo completan. Esta vertiente de la NEK europea tuvo como principal objeto de estudio el desempleo involuntario y el poder de negociación, característico del mercado laboral europeo de los años setenta.

[8] Entre las causas de esta asimetría de poder se encuentran: 1. el escaso atractivo de los desempleados de larga duración para las empresas (por su descapitalización); 2. los elevados costes de rotación o de ajuste, que merman la capacidad de los *outsiders* de competir por los puestos cubiertos; 3. las rentas situacionales (los *insiders* pueden contar con un elevado *know-how* o capital humano específico de la empresa); y 4. los factores institucionales (p. ej. la pertenencia a un sindicato o la legislación de protección del empleo de los *insiders*).

[9] La exposición canónica del modelo es la de Layard, Nickell y Jackman, *Unemployment: Macroeconomic Performance and the Labour Market* (1991), que lo contrasta con datos de los países de la OCDE.

Idea: dos agentes con poder de mercado, los sindicatos en el mercado de trabajo y las empresas en el de bienes[10], se reparten la renta nacional negociando en términos reales[11], y el reparto del poder depende de factores institucionales[12].

Planteamiento: con rendimientos constantes y el trabajo como único factor variable ($Y = A_L L$), el modelo se resume en dos relaciones:

- *Salario real negociado* (SRN): lo fijan los sindicatos al comienzo del periodo en términos nominales, a partir del nivel de precios esperado, pues no conocen los precios efectivos. Depende negativamente del desempleo y positivamente de los factores institucionales[13], y presenta inercia nominal.

- *Salario real pagado* (SRP): resulta del margen que las empresas aplican sobre el coste laboral unitario, $w/P = A_L/(1+\mu)$. El SRP es independiente del nivel de empleo y disminuye con el poder de mercado[14] y las cotizaciones a cargo de la empresa.

Dado el salario nominal, las variaciones de demanda agregada no alteran los precios (la productividad aparente es constante), sino las existencias o la producción, lo que respalda el argumento keynesiano del ajuste vía cantidades.

Equilibrio: en el plano empleo-salario real (figura A7.1), en el punto en que SRN = SRP y $P = P^e$ se determinan el salario real, el empleo y un desempleo de equilibrio único y estable, la tasa de paro que no acelera la inflación (NAIRU). La NAIRU es compatible con el desempleo involuntario (el salario se fija por encima del de vaciado) y con el desempleo estructural y friccional[15].

[10] La participación respectiva de cada parte en la renta será $w/(P \cdot A_L) = 1/(1+\mu)$ en el caso de los trabajadores y $\mu\, w/(P \cdot A_L) = \mu/(1+\mu)$ en el de las empresas.

[11] En una economía abierta se negocia el salario real de consumo, deflactado por el IPC.

[12] Además, el reparto del poder varía según la clase de trabajo (p. ej. cualificado o no) y según el estado del mercado laboral (p. ej. en un auge hay poco paro y una huelga le cuesta más a la empresa, lo que refuerza la posición del sindicato).

[13] Formalmente, SRN $= w/P^e = f(u^-, \varepsilon^-_{L^d w}, z^+)$. La relación inversa respecto de $\varepsilon_{L^d w}$ se debe a que una mayor elasticidad facilita a la empresa la sustitución de trabajadores.

[14] El grado de competencia depende en gran medida de la regulación. Por ejemplo, la culminación del mercado interior europeo (Acta Única Europea, 1986; mercado interior en vigor desde el 1 de enero de 1993) eliminó barreras no arancelarias, elevó la competencia y obligó a las empresas a reducir sus márgenes.

[15] A partir del nivel de desempleo de equilibrio pueden deducirse otras dos variables. La primera es el nivel de empleo de la economía. Teniendo en cuenta que PA $= N_o + U$, donde PA es la población activa, N_o la ocupada y U los parados, el empleo será $N_o = (1-u)\,\text{PA}$, con $u = U/\text{PA}$ la tasa de desempleo. La segunda, a partir de la relación entre empleo y desempleo, es la relación entre producción y desempleo. Para una función de producción $Y = A_L\, L$ y considerando $L \equiv N_o$, se tiene $u = 1 - Y/(A_L \cdot \text{PA})$.

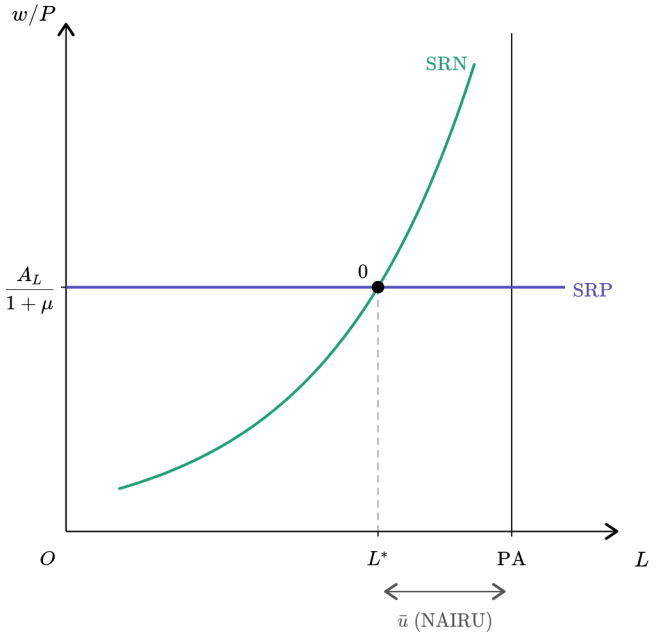

Figura A7.1: Determinación de la NAIRU (Layard y Nickell, 1985). En el plano empleo-salario real, el salario real negociado (SRN), creciente con el empleo y, por tanto, decreciente con el desempleo, corta al salario real pagado (SRP), horizontal en $A_L/(1+\mu)$, en el punto 0, con el empleo L^*. La distancia hasta la población activa (PA) es el desempleo correspondiente a la NAIRU, $\bar{u}$.

Estática comparativa: la NAIRU se desplaza con los factores institucionales (z) y el poder de mercado (μ):

- *Factores institucionales*: un aumento de z (p. ej. un mayor poder de negociación de los sindicatos) eleva el salario negociado para cada tasa de desempleo, desplaza la SRN hacia arriba y eleva la NAIRU.

- *Poder de mercado*: un aumento de μ (o de las cotizaciones a cargo de la empresa) reduce el salario real pagado, desplaza la SRP hacia abajo y eleva también la NAIRU.

- *Pendientes y elasticidades*: como la SRP no depende del empleo, cualquier desplazamiento mueve más la NAIRU cuanto más plana sea la SRN (cuanto menos responda el salario negociado al desempleo). Asimismo, una mayor elasticidad de la demanda de trabajo reduce el salario negociado y la NAIRU.

Extensiones: sin imponer $P = P^e$ se obtienen, además, una curva de Phillips de la NEK de primera generación (NEK1G) y una función de oferta agregada creciente en el nivel de precios, de las que se derivan las implicaciones de política del apartado 3.3. Esa curva es la tradicional ampliada por las expectativas, aquí adaptativas, $\pi_t = \pi_t^e - \kappa_u(u_t - \bar{u})$.

Conclusiones:

- *Aspectos positivos*: fundamenta el desempleo de equilibrio en factores institucionales y en el poder de mercado de sindicatos y empresas, lo que lo hace más rico que la tasa natural friedmaniana y permite analizar las políticas estructurales.

- *Limitaciones*: la NAIRU no es observable y sus estimaciones son inestables. Además, el modelo es estático y no explica por sí solo la persistencia del desempleo europeo tras las perturbaciones de los años setenta, que motiva la hipótesis de histéresis.

> El desarrollo completo del modelo (planteamiento, equilibrio, estática comparativa y derivación de la curva de Phillips) se encuentra en el tema A28, apartado 2.2, y la oferta agregada, en el tema A32, apartado 2.2.3.

1.2.2. Rigideces reales en el mercado de bienes y servicios

La fundamentación microeconómica de las rigideces reales en el mercado de bienes se apoya en la información y la competencia imperfectas. En condiciones normales los costes marginales caen mucho en las recesiones, por lo que las empresas tendrían incentivos a bajar precios. Para que no lo hagan ha de concurrir alguno de dos factores reales, bien que los costes marginales no caigan, bien que la elasticidad de la demanda caiga mucho ante una perturbación negativa.

1.2.2.1. Mercados de clientela: contratos implícitos en el mercado de bienes (Okun, 1981)

Origen: Arthur Okun (1981) traslada al mercado de bienes la lógica de los contratos implícitos. Como buscar y cambiar de proveedor tiene un coste para el cliente, en numerosos mercados se forma entre este y la empresa un vínculo de confianza.

Idea: las empresas no suben su precio en las expansiones para conservar a sus clientes, de modo que el margen en competencia monopolística resulta anticíclico.

Razonamiento: los costes de cambio y de búsqueda fidelizan al cliente y hacen que la elasticidad-precio de la demanda sea menor (más rígida) a corto plazo. Además, la elasticidad es procíclica, pues en fases de auge esos costes disminuyen al aumentar la actividad y el número de intercambios. Como a largo plazo las relaciones se difuminan y la elasticidad aumenta, las decisiones de precios adquieren un carácter intertemporal, y las empresas sacrifican beneficios a corto plazo con tal de mantener la relación con sus clientes.

Conclusiones:

- *Aspectos positivos*: explica la estabilidad de los precios en mercados con relaciones duraderas entre empresa y cliente a partir de la reputación y de los costes de cambio, un rasgo que los modelos de vaciado instantáneo ignoran.

- *Limitaciones*: el argumento es sobre todo verbal. La relación de confianza no se formaliza y es difícil de observar, por lo que la teoría resulta difícil de contrastar.

1.2.2.2. Economías de aglomeración o *thick market externalities* (Diamond, 1982)

Origen: Peter Diamond (1982) estudia una economía con fricciones de búsqueda en la que cuantos más agentes participan en el mercado más fácil resulta encontrar con quién intercambiar.

Idea: estas externalidades de mercado denso hacen que el coste marginal tienda a caer en los auges (*booms*) y a subir en las recesiones, de modo que las empresas no bajan sus precios en las recesiones.

Razonamiento: pueden existir equilibrios de alta y baja actividad. En los auges caen los costes de adquisición de *inputs* (y los costes medios), y a la inversa en las recesiones. Por ello, dadas la competencia imperfecta y los fallos de coordinación, el precio deseado apenas cae cuando lo hace la demanda y se producen rigideces.

Conclusiones:

- *Aspectos positivos*: muestra que los fallos de coordinación pueden dejar a la economía atrapada en un equilibrio de baja actividad, lo que ofrece un fundamento a las políticas de demanda.

- *Limitaciones*: no determina cuál de los equilibrios se alcanza, que depende de las expectativas, y la relación entre la densidad del mercado y los costes es difícil de cuantificar.

1.2.2.3. Guerras de precios durante los *booms* (Rotemberg y Saloner, 1986)

Origen: Julio Rotemberg y Garth Saloner (1986) estudian un oligopolio que sostiene la colusión tácita con la amenaza de castigar las desviaciones, en un juego repetido con demanda fluctuante pero observable (tema A19, apartado 2.2).

Idea: en los auges aumentan los incentivos a romper la colusión, por lo que el precio colusivo sostenible y el *mark-up* son anticíclicos.

Razonamiento: en los auges (demanda corriente elevada) la ganancia inmediata de desviarse crece respecto del castigo futuro, que depende de la demanda esperada. Aumentan así los incentivos a romper la colusión y entrar en guerras de precios, y el precio colusivo sostenible ha de ser más bajo. Como el margen cae cuando aumenta la actividad, el precio deseado responde poco a la demanda, lo que constituye una rigidez real.

Extensiones: Rotemberg y Woodford (1992) llevan el mecanismo a un modelo de equilibrio general dinámico calibrado. En ese modelo, la respuesta de la economía a una perturbación de demanda (un aumento del gasto militar) se aproxima más a la estimada con datos de posguerra de EE. UU. que la de un modelo competitivo.

Conclusiones:

- *Aspectos positivos*: da un fundamento estratégico, basado en los juegos repetidos, al carácter anticíclico del margen.

- *Limitaciones*: el resultado exige que las perturbaciones de demanda sean observables y transitorias (si fueran persistentes, el castigo futuro también crecería en los auges). Con información imperfecta, Green y Porter (1984) obtienen el resultado contrario, y la evidencia sobre la ciclicidad del margen no es concluyente.

1.2.2.4. Costes marginales constantes o decrecientes (Stigler, 1939)

Origen: otra causa de rigidez de precios es que los costes marginales sean constantes o decrecientes en la producción, lo que puede deberse a la capacidad ociosa que las empresas mantienen como reserva (Stigler, 1939) para afrontar las vicisitudes económicas.

Idea: si el coste marginal apenas cae en las recesiones, tampoco lo hace el precio deseado, que es un margen sobre ese coste.

Razonamiento: las perturbaciones de demanda se absorben vía cantidades. Con un coste marginal plano, además, la ganancia de ajustar el precio nominal es pequeña, por lo que esta rigidez real amplifica la de los costes de menú.

Conclusiones:

- *Aspectos positivos*: es la explicación más sencilla de la rigidez del precio real y encaja con la capacidad ociosa que las empresas mantienen en las recesiones.

- *Limitaciones*: traslada la explicación a la forma de la función de costes, que se supone en lugar de derivarse, y no explica por qué las empresas mantienen esa capacidad.

1.2.3. Conclusiones

Aspectos positivos: estas primeras aportaciones ofrecieron una fundamentación microeconómica de la rigidez de salarios y precios reales a partir de la información asimétrica y la competencia imperfecta, sin recurrir a la irracionalidad de los agentes.

Limitaciones: como señalan Laurence Ball y David Romer (1990), la rigidez real no implica rigidez nominal. Sin una fuente independiente de rigidez nominal, los precios se ajustan por completo a los *shocks* nominales pese a existir rigideces reales, porque la rigidez real amplifica la rigidez nominal, pero no la genera. Los modelos de rigideces reales no bastaban, pues, para romper la dicotomía clásica, lo que situó en el centro del programa el estudio de las rigideces o fricciones nominales de salarios y precios.

1.3. Teorías sobre rigideces nominales

Tres aportaciones fundamentan la rigidez nominal: la cuasirracionalidad, los costes de menú y las externalidades de demanda agregada.

1.3.1. Rigidez por racionalidad incompleta o cuasirracionalidad (Akerlof y Yellen, 1985)

Origen: George Akerlof y Janet Yellen (1985) parten del concepto de racionalidad incompleta o *cuasirracionalidad* en un marco macroeconómico de competencia imperfecta, en el que una proporción de empresas actúa de forma limitadamente racional.

Idea: esta cuasirracionalidad lleva a las conocidas reglas prácticas (*rules of thumb*), según las cuales las empresas no ajustan sus precios, porque la pérdida de no hacerlo es de segundo orden (pequeña).

Razonamiento: la existencia de un pequeño número de empresas cuasirracionales que asumen pérdidas de segundo orden da lugar a una pérdida de primer orden en el agregado.

Conclusiones:

- *Aspectos positivos*: ofrece una explicación sencilla de la rigidez nominal que solo exige desviaciones mínimas de la racionalidad, y su argumento sirve también de base al modelo de costes de menú.

- *Limitaciones*: muchas otras políticas de precios (y no solo mantenerlos fijos) producen pérdidas pequeñas, de modo que la cuasirracionalidad no es condición suficiente para la rigidez nominal.

1.3.2. Rigidez por costes tecnológicos: los costes de menú (Mankiw, 1985)

Origen: Gregory Mankiw (1985) se apoya en los costes de menú y completa la aportación de Akerlof y Yellen.

Idea: a la empresa no le compensa cambiar el precio cuando la pérdida prevista por mantenerlo, ya de por sí pequeña, es menor que los costes de menú.

Razonamiento: por *costes de menú* se entiende cualquier coste, por lo común reducido, que soporta la empresa cuando cambia su precio:

1. *Costes administrativos*: físicos (p. ej. reimprimir las tarifas y repartirlas[16]) y de información (p. ej. reunir y analizar los datos con los que fijar el nuevo precio óptimo).

2. *Costes no administrativos o exógenos*: p. ej. daño a la imagen, marcha de clientes fieles o estallido de una guerra de precios.

Mankiw llega así a la conclusión de Akerlof y Yellen sin renunciar a la racionalidad de los agentes, pues la empresa maximiza beneficios dentro del margen que le dejan los costes de menú.

Desarrollo gráfico: el modelo de costes de menú se representa en el plano producción-precio (figura A7.2):

1. *Curvas de demanda*, $D_0(x)$ y $D_1(x)$: son las demandas del competidor monopolista antes y después de la perturbación, decrecientes porque la empresa solo vende más si baja su precio.

2. *Curvas de ingreso marginal*, $\text{Img}_0(x)$ e $\text{Img}_1(x)$: quedan por debajo de la demanda correspondiente, porque vender una unidad más exige bajar el precio de todas.

3. *Coste marginal*, $\text{Cmg}(x)$: es creciente (rendimientos decrecientes).

4. *Equilibrio*: el del competidor monopolista, en el punto 0 (x_0, p_0). La producción x_0 iguala $\text{Img}_0(x)$ y $\text{Cmg}(x)$, y el precio p_0 se lee sobre $D_0(x)$.

[16] En hiperinflación o con gran volatilidad de precios, reajustarlos constantemente puede suponer un coste muy importante.

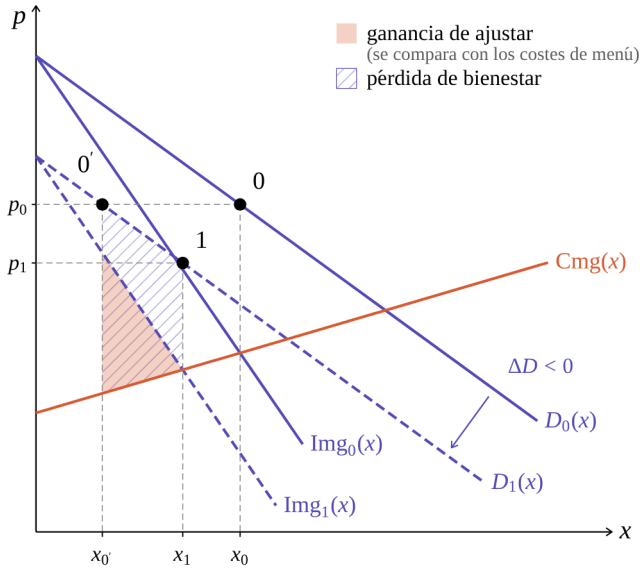

Figura A7.2: Costes de menú (Mankiw, 1985). Ante una caída de la demanda de D_0 a D_1, la empresa puede mantener el precio (punto $0'$) o ajustarlo (punto 1). Solo lo cambia si el triángulo de beneficios del ajuste supera los costes de menú. Si no lo hace, la producción queda por debajo de la óptima y se produce una pérdida de bienestar.

Estática comparativa: una perturbación negativa ($\Delta D < 0$) reduce la demanda de $D_0(x)$ a $D_1(x)$. La empresa puede:

- *No cambiar el precio* (punto $0'$, $x_{0'}, p_0$): ajuste íntegro vía cantidades, con ingreso marginal superior al coste marginal.

- *Cambiar el precio* (punto 1, x_1, p_1): ajuste vía precios y cantidades, igualando ingreso y coste marginal y optimizando.

El triángulo de beneficios de cambiar el precio (la *ganancia de ajustar*) determina la decisión, ya que la empresa solo lo modifica si ese beneficio supera los costes de menú. Si los precios permanecen estáticos, se produce una pérdida de bienestar social, pues la cantidad es menor que la óptima.

Extensiones (factores que afectan a la decisión):

1. *Estructura de costes* (tecnología): cuanto más inclinada sea la curva de coste marginal en el plano (x, p), mayor es el triángulo de beneficios y mayores los incentivos a cambiar el precio, mientras que una curva más plana (mayor rigidez del salario real) reduce esos incentivos.

2. *Estructura del ingreso* (poder de mercado): la pendiente de la curva de ingreso marginal, reflejo del poder de mercado de la empresa, condiciona también el tamaño del triángulo de beneficios y, con él, el incentivo a cambiar el precio.

Así, factores estructurales o reales (tecnología, estructura de mercado, horizonte temporal) pueden acentuar o atenuar la rigidez nominal que generan unas fricciones nominales dadas.

Conclusiones:

▪ *Aspectos positivos*: obtiene la rigidez nominal con empresas plenamente racionales, de modo que fricciones pequeñas bastan para que las perturbaciones de demanda tengan efectos reales y pérdidas de bienestar. Ball, Mankiw y Romer (1988) aportaron, además, un contraste frente a Lucas (1973). En los países con más inflación media, las empresas revisan sus precios más a menudo y la curva de Phillips a corto plazo es más vertical, como predicen los costes de menú, mientras que la varianza de la demanda, que destaca Lucas, explica menos.

▪ *Limitaciones*:

1. *Análisis dinámico*: el modelo es estático. Los costes de menú se pagan una sola vez, pero, si la perturbación es permanente, las pérdidas por no ajustarse se acumulan. Las empresas tendrán entonces más incentivos a cambiar el precio y el ajuste será más rápido (tanto más cuanto más largo sea el horizonte temporal).

2. *Equilibrios múltiples*: la teoría admite varios equilibrios estables, que antes no eran posibles.

3. *Costes de ajuste de la producción*: el modelo no considera los de las cantidades, que pueden ser mayores que los de los precios.

1.3.3. Externalidades de las fricciones nominales (Blanchard y Kiyotaki, 1987)

Origen: Blanchard y Kiyotaki (1987) aceptan tanto la cuasirracionalidad como los costes de menú y amplían el análisis al equilibrio general con competencia monopolística, enfatizando las externalidades de demanda.

Idea: parten de una función de demanda de la empresa que depende del *stock* real de dinero y del precio relativo de su bien[17], de modo que cambios en el *stock* real de dinero desplazan dicha demanda.

Razonamiento: el precio de cada empresa afecta, vía el nivel de precios y el *stock* real de dinero, a la demanda de las demás (externalidad de demanda agregada que ninguna internaliza). Así, las rigideces nominales individuales se acumulan y producen una inercia nominal de precios que hace fluctuar la producción mientras el nivel de precios permanece constante.

Conclusiones:

▪ *Aspectos positivos*: muestra que, en competencia monopolística, la producción de equilibrio es inferior a la eficiente y que, por la externalidad de demanda agregada, una expansión monetaria con precios rígidos puede mejorar el bienestar de todos los agentes. Da así a la política de demanda un fundamento microeconómico.

▪ *Limitaciones*: las mismas que las de la cuasirracionalidad y los costes de menú (apartados 1.3.1 y 1.3.2).

[17] Formalmente, se apoyan en una función del tipo $y_i^d = (M/P)\,(p_i/P)^{-\sigma}$, donde σ refleja el grado de sustituibilidad del bien.

1.3.4. Conclusiones

Aspectos positivos: el estudio de las fricciones nominales dio un fundamento microeconómico a la rigidez nominal y justificó la no neutralidad, yendo un paso más allá del modelo de la síntesis neoclásica (SNC).

Limitaciones: las fricciones nominales en competencia imperfecta no bastan para explicar el tamaño y la persistencia de los efectos reales de los *shocks* monetarios. La duda sobre si podían ser lo bastante importantes para justificar el no ajuste de precios abrió un programa de investigación paralelo sobre cómo se fijan los precios.

1.4. Análisis dinámico de las rigideces

El estudio dinámico de cómo se fijan los precios combina las rigideces nominales con el poder de mercado y se plantea dos preguntas, cada cuánto se revisan los precios y si las empresas los revisan a la vez o en momentos distintos.

1. *Sincronía*: cuando todas revisan al mismo tiempo, los *shocks* solo tienen efectos transitorios, y fricciones nominales como los costes de menú no explican por qué sus efectos reales persisten.

2. *Solapamiento*: cuando las revisiones se escalonan, los efectos reales duran mucho más. Este solapamiento puede obedecer a que las empresas son heterogéneas o a ventajas informacionales[18].

Los modelos de fijación de precios se clasifican en dos tipos, la fijación *en función del tiempo* (la revisión se realiza con una frecuencia dada) y la *dependiente del estado* (la revisión se desencadena por un suceso).

1.4.1. Reglas de ajuste en función del tiempo (Fischer, 1977; Taylor, 1980; Calvo, 1983)

Origen: la fijación escalonada de precios en función del tiempo parte de dos trabajos previos:

- *Fijación escalonada basada en la información disponible* (Fischer, 1977): Stanley Fischer (1977) considera agentes que deciden en $t-j$ sobre precios o salarios de t con la información disponible. Si no aciertan, el mercado no se vacía (a diferencia de Lucas, que supone vaciado) y la demanda determina las transacciones. Incluso la política monetaria sistemática tiene así efectos reales a muy corto plazo, pues puede reaccionar a información posterior a la firma de los contratos. El modelo no explica, sin embargo, la persistencia.

- *Contratos escalonados* (Taylor, 1980): John Taylor (1980) da el paso siguiente con los contratos escalonados, en los que los salarios se fijan por periodos fijos y solapados entre empresas. De este modo, el nivel de precios hereda inercia y las perturbaciones nominales tienen efectos reales persistentes, lo que convierte al modelo de Taylor en el antecedente directo del de Calvo.

Sobre esa base, Guillermo Calvo (1983)[19] formaliza la fijación escalonada de precios sobre la que se microfundamenta el modelo canónico de la NEK.

[18] Intuitivamente, las empresas tienen incentivos a esperar a que sus competidores fijen precios para recopilar información, de modo que se rompe la sincronización.

[19] Con todo, una vez agregado y linealizado, el modelo de costes cuadráticos de ajuste de Julio Rotemberg (1982) desemboca en la misma relación, y el de contratos escalonados de John Taylor (1980), en una de forma parecida, con más inercia (Roberts, 1995).

Idea: en cada periodo solo una fracción de las empresas puede revisar su precio, de modo que el nivel de precios se ajusta gradualmente y las perturbaciones nominales tienen efectos reales persistentes. La curva de Phillips de la NEK se deriva de este mecanismo.

Supuestos:

▪ *Entorno*: marco dinámico y estocástico con horizonte infinito y competencia monopolística en el mercado de bienes. En cada periodo, cada empresa recibe "luz roja" con probabilidad ξ (mantiene el precio) o "luz verde" con probabilidad $1-\xi$ (lo revisa), y ξ (constante y exógeno) mide así el grado de rigidez nominal[20]. La revisión es independiente del tiempo transcurrido y de la desviación respecto del precio óptimo, lo que le confiere naturaleza estocástica[21]. El índice general de precios es un promedio ponderado por la rigidez de los precios no revisados, P_{t-1}, y de los revisados, P_t^*:

$$P_t = P_{t-1}^{\xi}\, P_t^{*\,(1-\xi)} \;\rightarrow\; p_t = \xi\, p_{t-1} + (1-\xi)\, p_t^*$$

▪ *Agentes*: numerosas empresas y consumidores que forman expectativas según la HER y aplican un factor de descuento constante y exógeno $\beta \in (0,1)$. Los costes marginales son simétricos entre empresas y los consumidores tienen preferencia por la variedad.

▪ *Tecnología*: el coste marginal es procíclico con el PIB, es decir, se mueve con las desviaciones cíclicas del *output* respecto de su nivel natural[22], por lo que, con precios rígidos, el margen es contracíclico. La prociclicidad del coste marginal puede deberse a la escasez de *inputs* en las expansiones, a rendimientos decrecientes o a la desutilidad marginal creciente del trabajo.

Planteamiento (fijación de precios): sin rigideces, en un contexto estático y de competencia monopolística, la empresa fijaría un margen constante sobre el coste marginal[23], $P_{i,t} = (1+\mu)\,\mathrm{Cmg}_t$. En un marco estocástico y dinámico, además tiene en cuenta la rigidez y el descuento, maximizando el valor actual esperado de sus beneficios, dado el índice de rigidez ξ, un factor de descuento estocástico $\gamma_{t,t+j}$ y su restricción de demanda[24]:

$$\max_{P_t^*} V_t^e = \sum_{j=0}^{\infty} \xi^j\, \mathbb{E}_t\left[\gamma_{t,t+j}(P_t^*\, Y_{t+j} - \mathrm{CT}_{t+j}(Y_{t+j}))\right], \qquad \xi \in [0,1)$$
$$\text{s.a.: } Y_{t+j} = (P_t^*/P_{t+j})^{-\sigma} C_{t+j}$$

CPO: simplificada, $p_t^* = f(\xi, \gamma, \mu^+, (\mathbb{E}_t[\pi_{t+j}])^+, (\mathbb{E}_t[\mathrm{Cmg}_{t+j}])^+)$, donde el efecto de ξ es ambiguo (con inflación esperada positiva, una mayor rigidez eleva el precio de reajuste). El grado de rigidez ξ aparece elevado a j porque la empresa contempla la posibilidad de seguir sin revisar el precio en el futuro, de modo que el efecto del descuento y de la rigidez se diluye en el tiempo[25].

[20] Los microdatos permiten calibrar ξ. Bils y Klenow (2004) hallaron que la frecuencia mediana de cambio de los precios de consumo de EE. UU. (un 21 % al mes) implicaba una duración mediana de 4,3 meses, pero, sin las rebajas temporales, la duración mediana sube a entre 8 y 11 meses (Nakamura y Steinsson, 2008). En la zona del euro, los precios duran de cuatro a cinco trimestres de media (Dhyne et al., 2006).

[21] Esto se diferencia del modelo de contratos escalonados de Taylor (1980), en el que las empresas que revisan sus precios en el periodo t están influidas por las que no los revisan (pero los revisaron en $t-1$) y esperan revisarlos en $t+1$. El precio de contrato mira así hacia atrás y hacia delante, $p_t^* = \frac{1}{2}\,(p_{t-1}^* + \mathbb{E}_t[p_{t+1}^*]) + \psi\, \tilde{y}_t$, con $\psi > 0$, donde $\tilde{y}_t$ recoge la presión de la demanda (brecha de producción).

[22] Formalmente, $\widehat{\mathrm{cmg}}_t = \eta\, \tilde{y}_t$, con $\eta > 0$. La derivación y la lectura de la prociclicidad se encuentran en el Anexo, (8).

[23] Con una demanda de variedades con elasticidad de sustitución constante (CES), $\sigma > 1$, el margen bruto es $1+\mu = \sigma/(\sigma-1)$ (véase la demanda de cada variedad en el Anexo, A.1).

[24] Proviene de la condición de equilibrio en el mercado de bienes y servicios y del programa de maximización de los hogares.

[25] En $j = 0$ la empresa acaba de revisar el precio ($\xi^0\, \gamma_{t,t} = 1$). Para $j > 0$, pondera cada periodo por la probabilidad ξ^j de seguir con el mismo precio. Con un factor de descuento constante β, $\gamma_{t,t+j} = \beta^j$ y el peso de cada periodo futuro es $(\beta\xi)^j$.

La intuición es que las empresas fijan su margen deseado (sin rigideces) sobre un promedio ponderado de su coste marginal corriente y esperado, teniendo en cuenta la inflación esperada, la rigidez y su impaciencia.

CSO: el valor esperado es estrictamente cuasicóncavo en P_t^* (y cóncavo en un entorno del óptimo) porque la demanda es isoelástica con $\sigma > 1$ y el coste marginal es no decreciente, de modo que la CPO caracteriza el máximo.

Solución: con elevada rigidez de precios, estos tardan en moverse hacia su nuevo nivel deseado. Al igual que el de Layard-Nickell, el modelo de Calvo permite construir la curva de Phillips de la NEK *à la* Calvo, que admite tres especificaciones:

▪ *En función del coste marginal*: se obtiene[26] una primera especificación en la que la inflación depende de la inflación esperada (HER) y de la desviación del coste marginal real respecto de su nivel sin fricciones[27]:

$$\pi_t = \beta\, \mathbb{E}_t[\pi_{t+1}] + \kappa_{\text{cm}}\, \widehat{\text{cmg}}_t$$
$$\kappa_{\text{cm}} = \frac{(1-\xi)(1-\beta\xi)}{\xi} \cdot \frac{1-\alpha}{1-\alpha+\alpha\sigma} \;\rightarrow\; \kappa_{\text{cm}} = f(\xi^-, \alpha^-, \sigma^-, \beta^-)$$

El parámetro $\kappa_{\text{cm}} > 0$ (la pendiente de la curva) decrece con el grado de rigidez, la participación del capital, la elasticidad de sustitución entre variedades y el factor de descuento:

1. *Rigidez* (ξ): a mayor rigidez, menor respuesta de la inflación al coste marginal, pues menos empresas ajustan sus precios respecto del promedio anterior (ajustan márgenes en lugar de precios).

2. *Sustituibilidad entre variedades* (σ): con rendimientos decrecientes ($\alpha > 0$), a mayor sustituibilidad, más responde la producción de cada empresa (y, con ella, su coste marginal) a su precio relativo, lo que frena el reajuste y reduce la respuesta de la inflación al coste marginal. En cambio, con $\alpha = 0$, σ no afecta a la pendiente.

3. *Participación del capital* (α, que en el modelo mide los rendimientos decrecientes del trabajo): el coste marginal de cada empresa crece con su propia producción. En consecuencia, la empresa que eleva su precio vende menos y reduce su coste marginal, lo que modera el reajuste y la respuesta de la inflación al coste marginal medio.

▪ *En función de la brecha de producción*: el problema de la especificación anterior es que el coste marginal no es observable[28], por lo que suele usarse una basada en la brecha de producción (*output gap*). Dado que el coste es procíclico, existe proporcionalidad entre la brecha de producción y la desviación del coste marginal ($\widehat{\text{cmg}}_t = \eta\, \tilde{y}_t$), de modo que:

$$\pi_t = \beta\, \mathbb{E}_t[\pi_{t+1}] + \kappa\, \tilde{y}_t, \quad \kappa = \eta\, \kappa_{\text{cm}} = \kappa_{\text{cm}} \left(\theta + \frac{\varphi+\alpha}{1-\alpha}\right) > 0, \qquad \kappa > 0,\ \beta \in (0,1)$$

Aquí θ es el coeficiente de aversión relativa al riesgo y φ, la inversa de la elasticidad de Frisch (Anexo, (8)). Cuando la producción supera su nivel potencial ($\tilde{y}_t = y_t - \bar{y}_t > 0$, es decir, brecha positiva), se genera inflación[29].

[26] Que la inflación no aparezca en desviaciones y el coste marginal sí se debe a una normalización según la cual el valor de estado estacionario de la inflación, correspondiente a la NAIRU o a la tasa natural de paro, es nulo.

[27] Intuitivamente, el coste marginal real (en logaritmos) sería $\ln \text{Cmg}_t - p_t$ (el coste marginal nominal deflactado), y el coste marginal real sin fricciones, $-\mu \approx \ln \text{Cmg}_t - p_t^*$.

[28] Las cuentas nacionales informan de los costes (p. ej. los salarios), pero no del coste marginal. Por ello, los costes laborales suelen emplearse como aproximación del coste marginal.

[29] La magnitud de la incidencia de la brecha de producción sobre la inflación dependerá, de nuevo, de la calibración del modelo. Siguiendo a Jordi Galí (2008), una especificación típica para datos trimestrales sería: 1. $\xi = 2/3$,

- *En función del desempleo cíclico*: tomando la desviación del desempleo respecto de la NAIRU:

$$\pi_t = \beta\,\mathbb{E}_t[\pi_{t+1}] - \kappa_u\,(u_t - \bar{u}), \qquad \kappa_u > 0,\ \beta \in (0,1)$$

Cuando el desempleo cae por debajo de la NAIRU ($u_t - \bar{u} < 0$), la inflación corriente se sitúa por encima de la esperada descontada ($\pi_t > \beta\,\mathbb{E}_t[\pi_{t+1}]$).

Conclusiones:

- *Aspectos positivos*: en los tres casos, la curva de Phillips neokeynesiana de segunda generación (CP-NEK 2G) se diferencia de la de primera generación por su naturaleza *forward-looking*, al incorporar el operador de expectativas en el término de inflación futura.

- *Limitaciones*: precisamente por ello, la CP-NEK 2G no logra explicar la persistencia de la inflación, lo que le ha valido críticas (apartado 3.4.1).

1.4.2. Reglas de ajuste en función del estado (Caplin y Spulber, 1987; Blanchard y Fischer, 1989)

Origen: Andrew Caplin y Daniel Spulber (1987) consideran reglas dependientes del estado, en las que la revisión de precios no sigue una frecuencia definida, sino que se produce cuando la desviación entre el precio vigente y el óptimo rebasa ciertos umbrales o bandas (modelos *Ss*).

Idea: la decisión de revisar el precio sigue una lógica de coste-beneficio explícita, en la que la empresa paga el coste de menú solo cuando la pérdida acumulada por no ajustar lo justifica.

Razonamiento: Caplin y Spulber muestran que, si la distribución de precios dentro de la banda es uniforme y la inflación es positiva, las empresas que ajustan en cada momento son exactamente las necesarias para que el nivel general de precios se mueva uno a uno con el dinero. Así, la rigidez individual no genera rigidez agregada y el dinero es neutral.

Conclusiones:

- *Aspectos positivos*: advierte de que la rigidez individual no implica por sí sola rigidez agregada, porque el resultado depende de qué empresas ajustan y cuándo. Además, el modelo hace endógena la frecuencia de ajuste, que en las reglas en función del tiempo es exógena.

- *Limitaciones*: Blanchard y Fischer (1989) matizan que la neutralidad depende de supuestos muy restrictivos (perturbaciones solo positivas y distribución uniforme inicial). Con *shocks* de ambos signos, las reglas dependientes del estado recuperan efectos reales del dinero, si bien menores que las reglas en función del tiempo[30].

que implica que los precios se mantienen, en promedio, tres trimestres; 2. $\beta = 0{,}99$, que implica una tasa de retorno de los activos del 4 %; 3. $\theta = 1$, que implica una función de utilidad intertemporal logarítmica; 4. $\varphi = 1$, que implica una elasticidad de Frisch unitaria; y 5. $\sigma = 6$, que implica un *mark-up* del 20 % ($1+\mu = \sigma/(\sigma-1)$).

[30] Golosov y Lucas (2007) reavivaron el resultado de neutralidad. En su modelo de costes de menú, calibrado con microdatos, ajustan las empresas cuyo precio está más alejado del óptimo (*efecto selección*), y el dinero resulta casi neutral. Los trabajos posteriores que debilitan ese efecto recuperan efectos reales apreciables.

1.5. Conclusiones

Aspectos positivos: Laurence Ball y David Romer (1990)[31] demostraron que la interacción entre rigideces reales y fricciones nominales en los mercados de bienes y de trabajo puede generar rigideces importantes para entender los ciclos. En presencia de estas características, los *shocks* nominales pueden tener efectos reales tanto mayores cuanto mayor sea el grado de rigidez, lo que permite romper la dicotomía clásica.

Limitaciones: la primera oleada de neokeynesianos no logró frenar el desarrollo de la nueva macroeconomía clásica (NMC). Michel De Vroey (2016) atribuye esto a la postura defensiva de la NEK1G, anclada en una metodología que se estaba quedando obsoleta. Por ello, una segunda generación aceptará las reglas de juego de la teoría del ciclo real, pero cambiará la esencia del modelo incorporando las aportaciones de sus compañeros de primera generación.

[31] De hecho, Mankiw y Romer (1991) identifican la interacción entre las imperfecciones nominales y reales como uno de los rasgos distintivos de la NEK.

2. Aplicaciones y desarrollos: el estudio del ciclo con rigideces

Actualmente, las aportaciones de la NEK se condensan en un modelo DSGE que permite estudiar los ciclos económicos, la política económica y la interrelación entre múltiples variables en un marco de rigideces.

2.1. Modelo de la NEK en economía cerrada: IS-PC-MR

Origen: a finales de los años noventa, Richard Clarida, Jordi Galí, Mark Gertler y Michael Woodford publican una serie de trabajos[32] que dan lugar a lo que hoy se denomina *modelo canónico de la NEK*. Aquí se presenta una versión simplificada, con la exposición gráfica de Carlin y Soskice (2005), que dieron al modelo el nombre de IS-PC-MR.

Idea: el modelo busca sintetizar el comportamiento de la economía en tres ecuaciones. Sobre la metodología del ciclo real, esas ecuaciones incorporan la competencia monopolística y la rigidez de precios *à la* Calvo. Así, la política monetaria tiene efectos reales a corto plazo, pero no a largo plazo.

2.1.1. Supuestos

Entorno: existe competencia monopolística en el mercado de bienes y servicios (por lo que las empresas son precio-determinantes) y rigideces nominales en el mercado de bienes finales (precios *à la* Calvo), mientras que el mercado de trabajo es competitivo.

Agentes: se acepta la HER y hay cuatro agentes: hogares (consumidores y oferentes de trabajo, aversos al riesgo), empresas (productoras en competencia monopolística), sector público y banco central.

Tecnología: por simplicidad, se prescinde del capital como factor productivo y, con ello, de la inversión privada.

2.1.2. Modelo

Planteamiento: el modelo consta de tres ecuaciones: la curva IS dinámica (mercado de bienes y servicios), la curva de Phillips de la NEK 2G (fijación de precios) y la regla de política monetaria (autoridad monetaria).

Curva IS dinámica: se obtiene a partir de las decisiones intertemporales de ahorro y consumo de los hogares (en concreto, de la ecuación de Euler y de la condición de equilibrio en el mercado de bienes) tras log-linealizar en torno al estado estacionario.

En niveles (con las variables en logaritmos, en minúsculas), la curva IS dinámica puede expresarse como:

$$y_t = \mathbb{E}_t[c_{t+1}] - \frac{1}{\theta}(i_t - \mathbb{E}_t[\pi_{t+1}] - \rho) + g_t$$
$$\rightarrow \; y_t = f((\mathbb{E}_t[c_{t+1}])^+, (\mathbb{E}_t[y_{t+1}])^+, g_t^+, r_t^-)$$

donde:

[32] El modelo canónico se recoge en el monográfico de Jordi Galí (2008). Este se apoya en Richard Clarida, Jordi Galí y Mark Gertler (2000) y en Tack Yun (1996), además de las aportaciones de la NEK1G.

- y_t es el *output* de la economía y g_t el gasto público no productivo.

- $\mathbb{E}_t[c_{t+1}] = \mathbb{E}_t[y_{t+1} - g_{t+1}]$ es el consumo esperado (expectativa formada según la HER).

- $\theta > 0$ es el coeficiente de aversión relativa al riesgo de los hogares (inversa de la elasticidad de sustitución intertemporal), constante y exógeno.

- $i_t = -\ln P_B$ es el tipo de interés nominal a corto plazo, relacionado con el precio de los bonos.

- $\mathbb{E}_t[\pi_{t+1}]$ es la inflación esperada (HER), con $\mathbb{E}_t[\cdot]$ la esperanza condicionada a la información disponible en t.

- $\rho > 0$ es la tasa de descuento intertemporal de los hogares (constante y exógena), que refleja la impaciencia en el consumo.

La demanda agregada depende positivamente del consumo esperado (y, desglosándolo, del *output* esperado) y del gasto público corriente, y negativamente del tipo de interés real, $r_t = i_t - \mathbb{E}_t[\pi_{t+1}]$. La aversión al riesgo modula la sensibilidad de la demanda al tipo de interés ($1/\theta$), y cuanto mayor es el deseo de alisar el consumo, menos reacciona el gasto a los tipos. A diferencia de la IS estática, la IS dinámica recoge la naturaleza previsora de los agentes, con incidencia directa sobre la demanda agregada a través del efecto riqueza, la impaciencia y el alisamiento del consumo.

Alternativamente, la ecuación puede presentarse en términos de la brecha de producción, $\tilde{y}_t = y_t - \bar{y}_t$, de modo que el tipo de interés real efectivo aparezca en desviación respecto del tipo de interés real natural de la economía, $\bar{r}_t$.

En línea con la concepción wickselliana, el *tipo de interés real natural* es el determinado por factores reales que, en ausencia de *shocks* de costes, garantiza la estabilidad de precios. Es compatible con la neutralidad monetaria a largo plazo y con el *output* de precios flexibles y, en ausencia de *shocks* de costes y de rigidez salarial real, es eficiente desde el punto de vista normativo. Depende positivamente de la impaciencia y negativamente de los *shocks* tecnológicos transitorios, que reducen el crecimiento esperado del *output* natural. La aversión relativa al riesgo amplifica el efecto de ese crecimiento esperado, $\bar{r}_t = \rho + \theta\, \mathbb{E}_t[\Delta \bar{y}_{t+1}]$, por lo que solo eleva el tipo natural cuando se espera crecimiento: $\bar{r}_t = f(\rho^+, \theta, a^-)$.

Intuitivamente, si aumenta la impaciencia, los hogares consumen más y ahorran menos hoy, y la menor oferta de fondos prestables eleva el tipo de interés. Lo mismo ocurre, con crecimiento esperado, si aumenta la aversión al riesgo, que refuerza el deseo de alisar el consumo adelantándolo. Con los *shocks* tecnológicos ocurre lo contrario. Si aumenta la producción potencial, aumentan tanto el consumo como el ahorro, y el mayor volumen de fondos prestables hace caer el tipo de interés.

Formalmente, para el caso sin sector público, la IS dinámica en términos de la brecha de producción es[33]:

$$\tilde{y}_t = \mathbb{E}_t[\tilde{y}_{t+1}] - \frac{1}{\theta}(r_t - \bar{r}_t) \;\rightarrow\; \tilde{y}_t = f((\mathbb{E}_t[\tilde{y}_{t+1}])^+, r_t^-, \bar{r}_t^+)$$

Esta forma añade dos intuiciones:

1. *Brecha esperada*: a igualdad de tipo de interés, la expectativa de una brecha de producción mayor eleva ya hoy el consumo y el gasto agregado, y con ellos la brecha corriente.

[33] Aplicando recursividad hacia delante puede mostrarse que la brecha de producción es proporcional a la suma de las desviaciones corrientes y anticipadas del tipo de interés real efectivo respecto del natural: $\tilde{y}_t = -(1/\theta) \sum_{k=0}^{\infty} \mathbb{E}_t[r_{t+k} - \bar{r}_{t+k}]$.

2. *Desviación del tipo de interés real*: cuando el tipo de interés real efectivo se sitúa por encima del natural ($r_t > \bar{r}_t$), la brecha de producción se vuelve negativa (la actividad cae por debajo de su potencial), y a la inversa. La política monetaria afecta a la actividad real por este canal.

Curva de Phillips de la NEK 2G: se obtiene del mecanismo de fijación de precios *à la* Calvo (apartado 1.4.1). Relaciona la inflación con la inflación esperada y la brecha de producción, de modo que determina la inflación dada una senda de la brecha, que queda determinada a su vez por la senda de tipos y de gasto público:

$$\pi_t = \beta\, \mathbb{E}_t[\pi_{t+1}] + \kappa\, \tilde{y}_t \;\rightarrow\; \pi_t = f((\mathbb{E}_t[\pi_{t+1}])^+, \tilde{y}_t^+)$$

donde $\beta \in (0,1)$ es el factor de descuento de los hogares, $\mathbb{E}_t[\pi_{t+1}]$ es la inflación esperada (HER), κ recoge la velocidad de ajuste de la inflación y $\tilde{y}_t$ es la brecha de producción. La inflación depende positivamente de ambas variables, por las razones de ese mecanismo.

La curva recoge también el efecto del gasto público sobre la inflación, que pasa por la brecha de producción. El gasto mueve los dos términos de esa diferencia. Por un lado, altera la producción a través de la IS dinámica y, por otro, la producción potencial, porque, en el equilibrio con precios flexibles, su efecto riqueza cambia la oferta de trabajo.

Regla de política monetaria: la conducta del banco central puede resumirse en una regla de Taylor[34]:

$$i_t = \bar{i}_t + \alpha_\pi \tilde{\pi}_t + \alpha_y \tilde{y}_t \quad \text{o, en desviaciones,} \quad \tilde{i}_t = \alpha_\pi \tilde{\pi}_t + \alpha_y \tilde{y}_t, \qquad \alpha_\pi > 1,\ \alpha_y \geq 0$$

donde $\bar{i}_t = \bar{r}_t + \pi^T$ es el tipo de interés nominal neutral, $\tilde{\pi}_t = \pi_t - \pi^T$ la desviación de la inflación respecto del objetivo y $\tilde{i}_t = i_t - \bar{i}_t$. La intuición es que, si la inflación supera su objetivo, el tipo de interés nominal debe aumentar más que la inflación (es decir, $\alpha_\pi > 1$, el *principio de Taylor*)[35] para que suba el tipo real y se estabilice la economía, y a la inversa.

Con el intercepto $\bar{i}_t$, la regla traslada al tipo de interés cualquier cambio del tipo natural, de modo que las perturbaciones que solo mueven $\bar{r}_t$ quedan plenamente acomodadas. En la práctica, el banco central no observa $\bar{r}_t$ y usa una estimación que revisa con retraso o un intercepto constante, $\rho + \pi^T$ (Galí, 2008). En ese caso, un cambio del tipo natural abre una diferencia entre el tipo real y el natural que desplaza la MR, como muestran las perturbaciones analizadas a continuación y en el apartado 3.1.2.

Al sustituir esta regla en la IS dinámica se obtiene una relación negativa entre la brecha de producción y la desviación de la inflación respecto del objetivo, que desempeña el papel de la curva MR (regla monetaria). La versión óptima de la MR bajo discrecionalidad se deriva en el tema A36, apartado 2.4. La idea es que el compromiso del banco central con la estabilidad de precios se traduce, ante mayor inflación, en un menor *output*.

Desarrollo gráfico: el equilibrio del modelo se representa en los planos brecha de producción-desviación del tipo de interés real respecto del natural, arriba, y brecha de producción-inflación, abajo (figura A7.3):

[34] Esta no es la única especificación de la regla de Taylor. Existen formulaciones en función del desempleo cíclico ($i_t = \bar{i}_t + \alpha_\pi \tilde{\pi}_t - \alpha_u\,(u_t - \bar{u})$) y otras, como la de Clarida, Galí y Gertler (2000), que incluyen las expectativas de inflación en lugar de la inflación contemporánea.

[35] Formalmente, esto se relaciona con la condición de Blanchard-Kahn, que garantiza que el equilibrio sea estable y único.

1. *Curva IS dinámica*, IS: en el diagrama superior, recoge la relación negativa entre la brecha de producción, $\tilde{y}$, y la desviación del tipo de interés real respecto del natural, $r - \bar{r}$, porque un tipo real superior al natural frena el gasto.

2. *Curva de Phillips de la NEK 2G*, CP: en el diagrama inferior, refleja la fijación de precios *à la* Calvo y crece con la brecha de producción, que eleva el coste marginal.

3. *Curva MR*, MR: en el diagrama inferior, recoge la relación negativa entre la brecha de producción y la inflación que resulta de la regla de política, porque ante más inflación el banco central sube el tipo real.

4. *Equilibrio*: en el punto 0, la inflación está en el objetivo, π^T, y la producción en su nivel natural, con el tipo de interés real igual al natural.

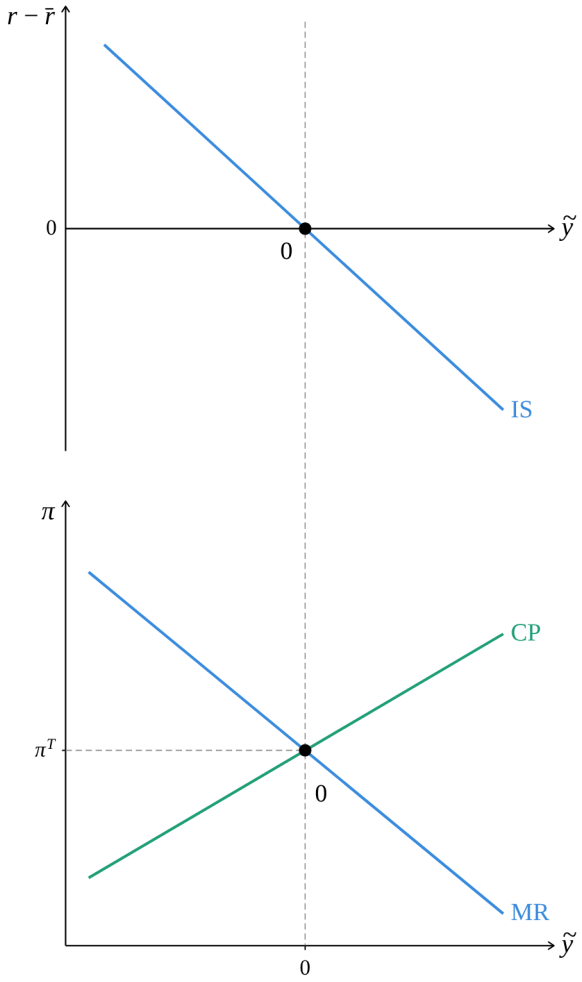

Figura A7.3: Equilibrio del modelo IS-PC-MR. La inflación está en el objetivo y la producción en su nivel natural (punto 0). Arriba, la IS dinámica en el plano brecha de producción-desviación del tipo de interés real respecto del natural, y abajo, la curva de Phillips de la NEK y la MR en el plano brecha de producción-inflación.

Estática comparativa: se consideran los desplazamientos de la CP (*shocks* de costes) y de la

MR (*shocks* que alteran el tipo de interés natural y se transmiten a través de la IS dinámica) y el ajuste a lo largo de ellas:

▪ *Perturbación tecnológica*: un *shock* tecnológico positivo eleva el *output* natural y reduce el tipo de interés natural. Si el banco central traslada plenamente esa caída a su tipo, la brecha de producción se cierra sin conflicto entre objetivos. Si no lo hace, p. ej. porque su regla usa una estimación del tipo natural que aún no recoge la caída, el tipo real queda por encima del natural y la MR se desplaza de MR_0 a MR_1. La economía pasa así del punto 0 al 1, con brechas negativas de producción e inflación (figura A7.4).

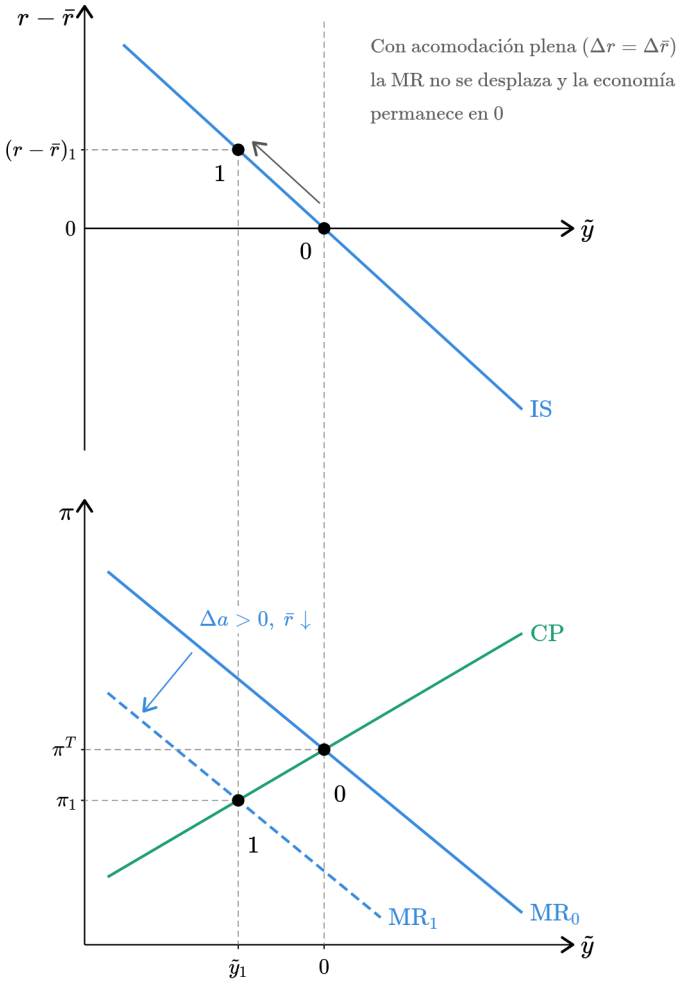

Figura A7.4: Modelo IS-PC-MR ante un *shock* tecnológico positivo ($\Delta a > 0$). Si la regla de Taylor no sigue la caída del tipo de interés natural, el tipo real queda por encima del natural, la MR se desplaza a la izquierda y la economía pasa de 0 a 1, con brechas de producción e inflación negativas. Con acomodación plena, la brecha no se abre y la economía permanece en 0.

▪ *Perturbación de costes*: un *shock* de costes ($\varepsilon^\pi > 0$) desplaza la CP hacia arriba, de CP_0 a CP_1 (figura A7.5). El *shock* obliga al banco central a aceptar transitoriamente una brecha negativa para devolver la inflación al objetivo. La economía se mueve así a lo largo de la MR del punto 0 al 1.

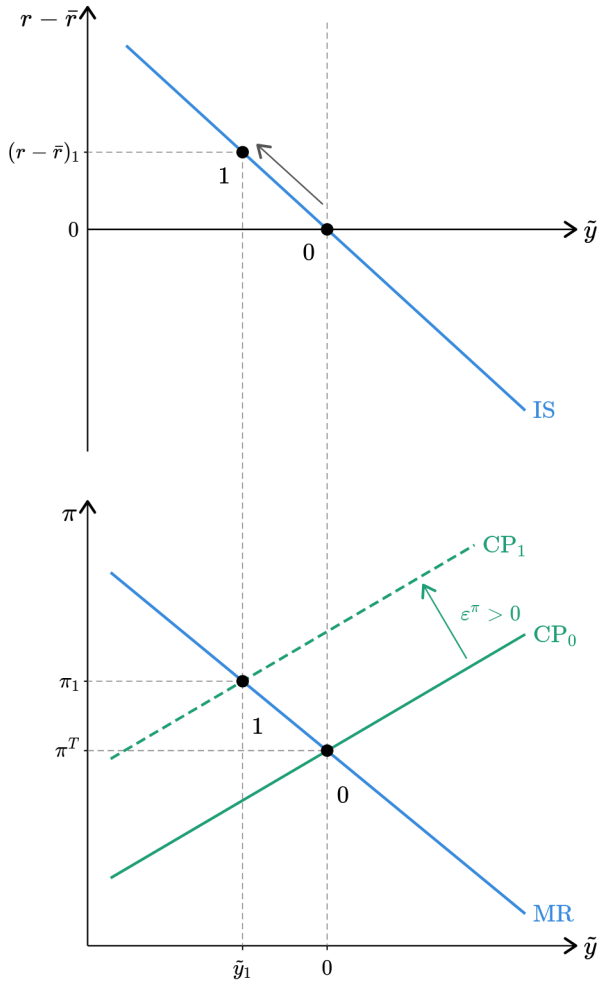

Figura A7.5: Modelo IS-PC-MR ante un *shock* de costes. El banco central acepta transitoriamente una brecha de producción negativa para devolver la inflación al objetivo.

2.1.3. Conclusiones

Aspectos positivos:

▪ *Microfundamentación*: el modelo dota a la macroeconomía keynesiana de una fundamentación microeconómica rigurosa (agentes optimizadores, ecuación de Euler), de modo que resulta, en principio, robusto a la crítica de Lucas.

▪ *Forma sintética*: sus tres ecuaciones (IS dinámica, CP-NEK y regla de política) son analíticamente tratables y permiten el análisis del bienestar. Por ello, el modelo se ha convertido en el de referencia (*workhorse*) para el estudio de la política monetaria y para los bancos centrales.

▪ *Estabilidad de precios*: ofrece un marco coherente para fundamentar ese objetivo y las reglas de tipos de interés (p. ej. la regla de Taylor y los objetivos directos de inflación o *inflation targeting*).

Limitaciones:

▪ *Persistencia de la inflación*: por su naturaleza *forward-looking*, la CP-NEK 2G no explica bien la persistencia o inercia de la inflación, lo que ha obligado a recurrir a versiones híbridas con indexación al pasado (apartado 3.4.1).

▪ *Agente representativo*: el modelo básico se apoya en un agente representativo y prescinde del capital, la inversión y el sector financiero, omisiones especialmente expuestas tras la crisis de 2008; tampoco recoge la heterogeneidad de los agentes ni el desempleo involuntario.

▪ *Supuestos exigentes*: las expectativas racionales y la log-linealización en torno al estado estacionario limitan su capacidad para tratar no linealidades, grandes perturbaciones o situaciones como el límite inferior de los tipos de interés (*zero lower bound*, ZLB).

2.2. Extensiones al modelo

El modelo base de tres ecuaciones admite extensiones que corrigen parte de las limitaciones anteriores.

2.2.1. Modelo de la NEK en economía abierta (Galí y Monacelli, 2005)

Origen: Jordi Galí y Tommaso Monacelli (2005) extienden el modelo a una economía abierta.

Idea: se mantiene la estructura de tres ecuaciones del modelo cerrado, pero cada una se modifica para recoger la apertura, y se añaden la paridad de intereses y la noción de tipo de cambio natural.

Supuestos: se mantienen los supuestos neokeynesianos de economía cerrada (competencia monopolística, rigideces nominales, HER, etc.) y se añade la apertura, en concreto, la de un país pequeño que interactúa con un único resto del mundo, sobre cuyas variables no influye. Se cumple la condición de Marshall-Lerner (las depreciaciones reales mejoran la balanza comercial) y los mercados financieros y de divisas son completos.

Planteamiento: con $p_{o,t}$ el precio relativo de los bienes importados, $\tilde{E}_t$ la desviación del tipo de cambio nominal respecto de su nivel natural y ε_t^d, ε_t^π y ε_t^m las perturbaciones de demanda, de costes y de política monetaria, las tres ecuaciones del modelo abierto son:

$$\begin{aligned}
\text{IS:}\ & \tilde{y}_t = \mathbb{E}_t[\tilde{y}_{t+1}] - \frac{1}{\theta_\alpha}(r_t - \bar{r}_t) + \Gamma_y\, p_{o,t} + \varepsilon_t^d, \qquad \Gamma_y \gtrless 0 \\
\text{CP:}\ & \pi_t = \beta\, \mathbb{E}_t[\pi_{t+1}] + \kappa_\alpha\, \tilde{y}_t + \Gamma_\pi\, p_{o,t} + \varepsilon_t^\pi, \qquad \Gamma_\pi > 0,\ \beta \in (0,1) \\
\text{Regla:}\ & i_t = \bar{i}_t + \alpha_\pi\, \tilde{\pi}_t + \alpha_y\, \tilde{y}_t + \alpha_E\, \tilde{E}_t + \varepsilon_t^m, \qquad \alpha_E \geq 0
\end{aligned}$$

1. *IS dinámica*: la sensibilidad al tipo de interés, $1/\theta_\alpha$, depende del grado de apertura $\alpha \in [0,1]$, el peso de los bienes importados en el consumo, porque los consumidores sustituyen entre bienes nacionales y extranjeros, no solo entre presente y futuro. Con $\alpha = 0$ se recupera el caso cerrado (no debe confundirse con la participación del capital del apartado 1.4.1). Al tratarse de un país pequeño, el tipo de interés real natural $\bar{r}_t$ depende también de la renta exterior, por lo que la economía responde a los *shocks* de productividad extranjeros. El signo de Γ_y es indeterminado *a priori*, porque un encarecimiento de los importados desvía demanda hacia los bienes nacionales (efecto sustitución) y, a la vez, reduce la renta real disponible (efecto renta).

2. *Curva de Phillips*: la inflación relevante es la del IPC, que incluye los bienes importados[36]. La apertura cambia la pendiente de la curva, κ_α, y el término $\Gamma_\pi\, p_{o,t}$ es el canal de la *inflación importada*, por el que una depreciación o una subida de los precios exteriores eleva la inflación aunque la brecha de producción no varíe.

3. *Regla de política*: el banco central puede responder, además, a la desviación del tipo de cambio respecto de su nivel natural, con flotación pura si $\alpha_E = 0$, flotación dirigida si $\alpha_E > 0$ y tipo fijo en el límite, $\alpha_E \to \infty$. Ante perturbaciones de demanda, oferta o inflación, parte del ajuste se canaliza a través del tipo de cambio real, de modo que el peso de la estabilización no recae ya exclusivamente en la autoridad monetaria y el tipo de interés.

Conclusiones:

▪ *Aspectos positivos*: el modelo ordena por su bienestar los regímenes monetarios de una economía abierta y muestra que, en un caso particular, estabilizar la inflación nacional es óptimo y la paridad fija, subóptima (apartado 3.1.4).

▪ *Limitaciones*: el supuesto de país pequeño excluye el efecto de la política nacional sobre el resto del mundo. Los mercados completos suponen, además, un reparto perfecto del riesgo entre países, que deja fuera las primas de riesgo y la dinámica de la deuda externa.

> El modelo de economía abierta, su equilibrio y los regímenes cambiarios se desarrollan en el tema A32, apartado 3.3.

2.2.2. Rigideces salariales reales y fricciones de búsqueda (Blanchard y Galí, 2007; 2010)

Olivier Blanchard y Jordi Galí completan el modelo en dos direcciones complementarias:

1. *Rigidez salarial real* (2007): los salarios reales se ajustan lentamente a las condiciones del mercado de trabajo. Esa rigidez genera en la CP-NEK un término de costes endógeno y, con él, una verdadera disyuntiva entre estabilizar la inflación y estabilizar la actividad (la ruptura de la "divina coincidencia", apartado 3.1.2).

2. *Fricciones de búsqueda* (2010): los autores incorporan fricciones de emparejamiento en el mercado de trabajo (*à la* Diamond-Mortensen-Pissarides), lo que dota al modelo de desempleo de equilibrio. Así, el marco neokeynesiano pasa a poder hablar explícitamente de las fluctuaciones del desempleo (ausentes en el modelo canónico, construido sobre un mercado de trabajo walrasiano).

> El modelo de búsqueda y emparejamiento se desarrolla en el tema A26, apartado 2, y su lugar en la teoría del desempleo de equilibrio, en el tema A28, apartado 3.1.

[36] En Galí y Monacelli (2005), la IS y la curva de Phillips se escriben en términos de la inflación interna (de los bienes nacionales) y no contienen términos en el precio de los importados. La inflación del IPC se obtiene después a partir de la interna y de la variación de la relación real de intercambio. Los términos en $p_{o,t}$ son una extensión estilizada, la misma del tema A32, apartado 3.3.1, para analizar perturbaciones en el precio de las importaciones, como las energéticas. Que estabilizar la inflación interna sea óptimo en el caso particular del apartado 3.1.4 se debe, precisamente, a que en el original la curva de Phillips está en esa inflación.

2.2.3. La curva de Phillips salarial y la NAWRU (Galí, 2011)

Jordi Galí (2011) recupera la curva de Phillips salarial dentro del marco neokeynesiano. Expresa la inflación salarial en función de la brecha de desempleo (la desviación del desempleo respecto de su tasa natural o NAWRU, *Non-Accelerating Wage Rate of Unemployment*):

$$\pi_t^w = \beta\, \mathbb{E}_t[\pi_{t+1}^w] - \kappa_w\, (u_t - \bar{u}_t)$$

Con ello reintroduce el desempleo como variable central y reconecta el modelo moderno con la relación de Phillips original entre salarios y desempleo.

> La curva de Phillips neokeynesiana, con la versión salarial de Galí (2011), y la estimación de la NAIRU se desarrollan en el tema A28, apartados 3.2 y 3.3.

2.3. Conclusiones

El modelo canónico condensa la aportación de la segunda generación neokeynesiana. Recupera la no neutralidad monetaria a corto plazo de la tradición keynesiana sin renunciar al rigor de los nuevos clásicos, y es por ello el punto de encuentro de la nueva síntesis neoclásica (apartado 3.5).

El modelo admite, además, extensiones como la economía abierta, las fricciones del mercado de trabajo, la heterogeneidad de los agentes o las fricciones financieras, que se incorporan a él, al menos en versiones simplificadas, sin alterar la estructura de tres ecuaciones.

3. Implicaciones del pensamiento de la NEK

Las implicaciones de la NEK se agrupan por ámbitos de política, y la sección se cierra con las críticas a la escuela y la nueva síntesis neoclásica.

3.1. Implicaciones de política monetaria

La principal conclusión de la NEK en materia de política monetaria es que esta puede tener efectos importantes y persistentes tanto sobre las variables reales como sobre las nominales de la economía.

3.1.1. La importancia de los niveles naturales de *output* y de tipo de interés

Los niveles naturales de *output* y de tipo de interés (los correspondientes al equilibrio con precios flexibles, es decir, sin rigideces nominales) desempeñan un papel central en el diseño de la política monetaria, pues constituyen un punto de referencia para los decisores. La política monetaria no puede mantener desviaciones persistentes respecto de ellos sin generar presiones inflacionarias o deflacionarias.

No obstante, como los niveles naturales no son observables, resulta complejo utilizarlos en la práctica. Por ello, tradicionalmente se aproximaban mediante tendencias estadísticas suaves (p. ej. filtros de series temporales), mientras que en los modelos actuales son los valores que se generarían en un modelo del ciclo real sin fricciones.

3.1.2. La divina coincidencia y la "no tan divina" coincidencia (Blanchard y Galí, 2007)

Olivier Blanchard y Jordi Galí (2007) acuñan el concepto de **"divina coincidencia"**, la posibilidad de estabilizar simultáneamente la inflación y la brecha del *output* respecto de su nivel eficiente con la política monetaria. En ausencia de imperfecciones reales no triviales (en concreto, de rigidez salarial real) no existe conflicto entre el objetivo de inflación y el de actividad, de modo que el banco central podría centrarse únicamente en la inflación y, con ello, estabilizar también la economía real.

Ahora bien, los propios autores reconocen que la inexistencia de una disyuntiva (*trade-off*) es poco verosímil empíricamente. Al ampliar el modelo con rigidez salarial real, la divina coincidencia se rompe y surge una disyuntiva entre ambos objetivos. La política monetaria óptima pasa entonces por tolerar algo más de inflación, ya que, de lo contrario, ciertos *shocks* adversos provocarían una caída del *output* ineficientemente elevada.

Otros autores rompen la divina coincidencia introduciendo una perturbación de costes (*cost-push*) en la CP-NEK, que recoge *shocks* exógenos por impuestos distorsionantes, salarios deseados o *mark-ups*:

$$\pi_t = \beta\, \mathbb{E}_t[\pi_{t+1}] + \kappa\, \tilde{y}_t + \varepsilon_t^{\pi}, \qquad \kappa > 0,\ \beta \in (0,1)$$

donde ε_t^{π} es el *shock* de oferta o de costes.

El componente de demanda de la COVID-19 se representa como una caída de la tasa de preferencia temporal ($\Delta\rho < 0$), en los planos brecha de producción-desviación del tipo de interés real respecto del natural y brecha de producción-inflación (figura A7.6). La caída recoge el mayor ahorro por precaución de los hogares y equivale a un mayor factor de descuento, $\beta = 1/(1+\rho)$, de modo que el tipo de interés natural cae. Como la regla de Taylor no sigue esa caída, el tipo

real queda por encima del natural y la MR se desplaza a la izquierda, de MR_0 a MR_1, y la economía pasa de 0 a 1, con brecha de producción e inflación negativas.

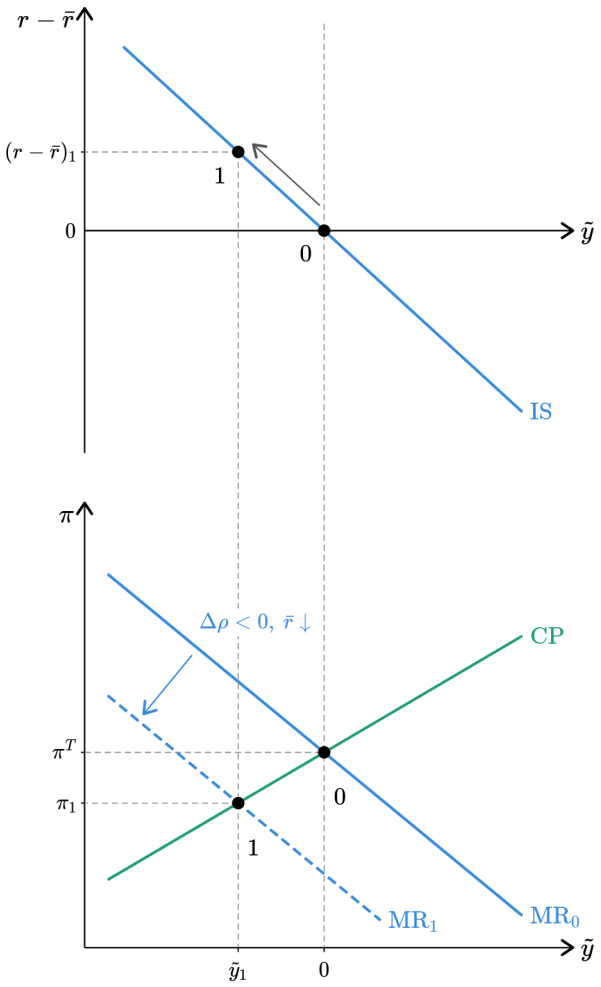

Figura A7.6: Modelo IS-PC-MR ante el componente de demanda del *shock* de la COVID-19. La caída de la tasa de preferencia temporal ($\Delta\rho < 0$), es decir, el mayor ahorro por precaución, reduce el tipo de interés natural. Como la regla de Taylor no sigue esa caída, la MR se desplaza a la izquierda, de MR_0 a MR_1, y la economía pasa de 0 a 1, con brecha de producción e inflación negativas.

3.1.3. La gestión de las expectativas: el *forward guidance* (Campbell et al., 2012)

Del modelo base se desprende la importancia de las expectativas en la transmisión de la política monetaria, ya que la inflación no solo es función de la actividad real (la brecha de producción), sino también de la senda anticipada de tipos de interés[37]. Por ello, la gestión de las expectativas del sector privado (y, en particular, la comunicación de las intenciones del banco central sobre los tipos) es determinante de la eficacia de la política monetaria.

A raíz de la Gran Recesión ha crecido el interés por el *forward guidance* (orientación prospectiva), práctica por la que el banco central comunica la senda futura de su política, como previsión

[37] Esto se deduce del hecho de que la brecha de producción depende de las expectativas sobre los tipos de interés.

(délfica) o como compromiso (odiseica, p. ej. estado-contingente). Esa comunicación ayuda a anclar las expectativas.

3.1.4. Estabilización de la inflación y *overshooting* (Dornbusch, 1976)

Del modelo de economía abierta se desprende que, bajo ciertas condiciones, la política monetaria óptima es la que estabiliza la inflación nacional y es, a la vez, acomodaticia con las fluctuaciones del tipo de cambio. Un problema asociado es la posibilidad de sobrerreacción del tipo de cambio (*overshooting*) a corto plazo.

Origen: el término lo acuña Rudiger Dornbusch (1976) en un modelo que combina el marco Mundell-Fleming con expectativas racionales y que anticipa la lógica neokeynesiana de precios rígidos y activos flexibles.

Idea: con precios de los bienes rígidos a corto plazo y un tipo de cambio flexible que se ajusta instantáneamente bajo expectativas racionales, el tipo de cambio sobrerreacciona ante una expansión monetaria. A corto plazo se deprecia más de lo que exige su nuevo equilibrio a largo plazo.

Razonamiento: una expansión monetaria no anticipada reduce el tipo de interés nacional porque los precios no se ajustan. Por la paridad descubierta de intereses, un tipo inferior al exterior exige que los agentes esperen una apreciación futura. Para que esa apreciación sea posible, el tipo de cambio debe depreciarse hoy por encima de su nivel de largo plazo (el que dicta la paridad del poder adquisitivo). Después, a medida que los precios se ajustan al alza, el tipo de interés recupera su nivel y el tipo de cambio se aprecia gradualmente hacia su equilibrio. El modelo se desarrolla en el tema B14, apartado 2.2.2.

Conclusiones:

- *Aspectos positivos*: la rigidez de precios, combinada con la flexibilidad de los mercados financieros, explica la elevada volatilidad del tipo de cambio sin recurrir a la irracionalidad de los agentes.

- *Limitaciones*: el tamaño de la sobrerreacción depende de parámetros difíciles de estimar, como la velocidad de ajuste de los precios, y la evidencia apunta a que la máxima depreciación llega con retraso respecto del *shock* monetario (*delayed overshooting*). Además, la sustituibilidad perfecta entre activos que exige la paridad descubierta de intereses excluye la prima de riesgo cambiario.

3.2. Implicaciones de política fiscal

Tres aportaciones analizan la eficacia y el diseño de la política fiscal en el marco neokeynesiano: las restricciones de liquidez, la respuesta de la política monetaria y la unión monetaria.

3.2.1. Efectos sobre el consumo: restricciones de liquidez (Galí, López-Salido y Vallés, 2007)

Jordi Galí, David López-Salido y Javier Vallés plantean las restricciones de liquidez como mecanismo que amplifica los efectos expansivos de la política fiscal. Parten del modelo canónico de la NEK con dos tipos de agentes:

1. *Agentes ricardianos*: no sufren restricciones de liquidez, participan en los mercados financieros y suavizan su consumo intertemporalmente (comportamiento racional estándar).

2. *Agentes no ricardianos* (*rule of thumb* o de "vivir al día"): están restringidos financieramente y consumen íntegramente su renta disponible corriente.

Dado que los agentes no ricardianos[38] no participan en los mercados financieros, no experimentan el efecto riqueza negativo asociado al aumento del gasto público (la anticipación de mayores impuestos futuros), lo que amplifica la eficacia de la política fiscal. De hecho, los autores muestran que, si existe una proporción suficiente de agentes no ricardianos, el multiplicador del gasto público puede ser superior a la unidad.

3.2.2. Relación con la política monetaria (Woodford, 2011; Eggertsson, 2011)

La respuesta de la política monetaria ante un *shock* fiscal altera el valor del multiplicador. Se distinguen tres escenarios:

1. *Tipo de interés real constante*: si el banco central mantiene constante el tipo de interés real, el multiplicador es igual a la unidad.

2. *Objetivo de inflación o regla de Taylor*: si la política monetaria persigue un objetivo estricto de inflación, el multiplicador es inferior a la unidad (la subida de tipos desplaza parcialmente al gasto privado).

3. *ZLB*: el multiplicador es mayor que la unidad cuando los tipos nominales están en cero, siempre que el banco central deje pasar el estímulo fiscal sin subirlos. La intuición es que, con el tipo nominal anclado en cero, el aumento del gasto eleva las expectativas de inflación y reduce el tipo de interés real.

Gráficamente, en los planos brecha de producción-desviación del tipo de interés real respecto del natural y brecha de producción-inflación (figura A7.7), la expansión ($\Delta g > 0$) eleva el tipo de interés natural y desplaza la MR de MR_0 a MR_1. A tipo real constante (primer escenario), la economía iría por la IS y la CP hasta $0'$. En cambio, con la regla de Taylor (segundo escenario), el banco central sube el tipo real ($\Delta r > 0$) y la economía queda en 1, con menos producción que en $0'$.

[38] Esta figura procede de Campbell y Mankiw (1989), que estimaron que en torno a la mitad de la renta de EE. UU. la reciben hogares que consumen su renta corriente.

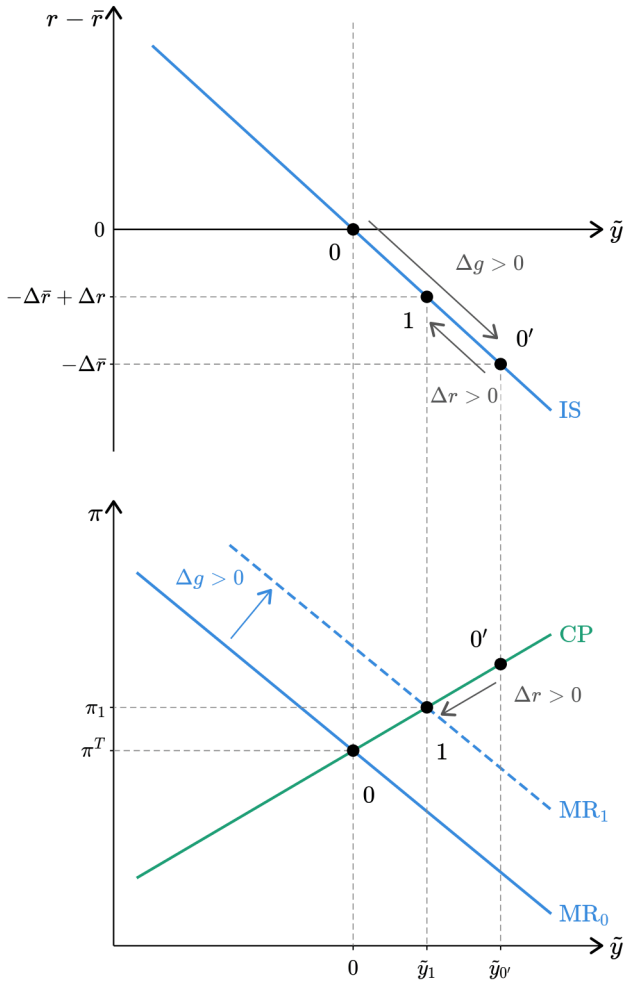

Figura A7.7: Modelo IS-PC-MR ante una expansión fiscal ($\Delta g > 0$). La expansión eleva el tipo de interés natural. A tipo real constante, la economía pasaría de 0 a $0'$ (multiplicador unitario), pero la respuesta del banco central según su regla ($\Delta r > 0$) la sitúa en 1, sobre la MR desplazada, con un multiplicador inferior a la unidad.

Por su parte, Eggertsson (2011) amplía el análisis incorporando perturbaciones en el factor de descuento de los hogares y los efectos de la política impositiva. Una perturbación en dicho factor lleva a los hogares a aumentar su ahorro precautorio, de modo que la restricción de tipos nominales nulos se vuelve vinculante. Con el *output* cayendo en un contexto recesivo y de rigideces, las expectativas de deflación elevan el tipo de interés real (el nominal no puede bajar de cero), lo que genera nuevos incentivos a ahorrar. La espiral de ahorro precautorio se traduce en una gran pérdida de *output*, y precisamente en ese contexto el multiplicador del gasto público resulta elevado.

3.2.3. Política fiscal y monetaria en una unión monetaria (Galí y Monacelli, 2008)

Jordi Galí y Tommaso Monacelli (2008) parten del modelo canónico de la NEK en economía abierta para analizar el diseño de la política fiscal y monetaria en una unión monetaria. En

ella, cada país conserva el control de su política fiscal (vía consumo público, es decir, gasto no productivo), mientras que la política monetaria es común[39].

La conclusión principal es que la política monetaria común debe estabilizar la inflación del conjunto de la unión. En cambio, la renuncia a una política monetaria propia, unida a la rigidez de precios, otorga a la política fiscal nacional un papel estabilizador ante perturbaciones específicas de cada país, más allá de la provisión eficiente de bienes públicos. Como el banco central común no responde a la inflación generada en un país pequeño concreto, sino a la del conjunto de la unión, el efecto del gasto público nacional no depende de la acomodación monetaria.

3.3. Implicaciones de política de empleo: políticas de demanda y estructurales (Layard y Nickell, 1985)

La política de empleo de la escuela se apoya en el modelo de negociación de Layard y Nickell (apartado 1.2.1.4). Ese modelo muestra que, al menos a corto plazo, las políticas de demanda recuperan cierto margen de actuación, si bien a costa de mayor inflación. Por ello, la escuela subraya las políticas estructurales y de rentas, orientadas a reducir la NAIRU desplazando las curvas SRN y SRP. Entre ellas destacan:

- *Cuña fiscal*: su reducción[40].

- *Formación*: el gasto público en formación, que eleva la productividad y reduce el poder de negociación de los trabajadores cualificados[41].

- *Subsidio de desempleo*: su reforma.

> Las políticas para reducir la NAIRU se analizan en detalle en el tema A28, apartado 2.2, y las políticas de empleo en general en el anexo del tema A27.

3.4. Críticas al pensamiento de la NEK

Las críticas a la NEK se dirigen a su curva de Phillips, a la ausencia del sistema financiero, al descuido de la política fiscal y a su metodología.

3.4.1. Críticas a la CP-NEK

La CP-NEK se enfrenta a tres críticas, todas relacionadas con su carácter prospectivo (*forward-looking*):

1. *Inmediatez del ajuste*: su carácter prospectivo predice un ajuste inmediato, cuando la evidencia muestra que el ajuste (y en especial la desinflación) es gradual (Mankiw, 2001).

2. *Ausencia de persistencia*: por el mismo motivo, no reproduce la inercia de la inflación.

[39] En términos del modelo canónico, supone que el tipo de interés de la IS dinámica de cada país es el fijado por el banco central común a todos los Estados miembros.

[40] Un aumento de los impuestos directos o de las cotizaciones sociales de los empleados reduce el salario neto que perciben los trabajadores, por lo que los sindicatos buscarán elevar el SRN para no perder poder adquisitivo, y la curva SRN se desplazará hacia arriba (una reducción tendría el efecto contrario). De forma análoga, una disminución de los impuestos indirectos o de las cotizaciones a cargo de los empleadores aumenta el SRP.

[41] Si las empresas tienen problemas para expandirse por falta de trabajadores cualificados, estos podrán elevar el SRN. La formación resuelve ese cuello de botella.

3. *Evidencia empírica*: en encuestas a más de 11.000 empresas de la zona del euro, la mayoría tiene en cuenta al fijar sus precios tanto la evolución económica pasada como la esperada (Fabiani et al., 2006), mientras que la CP-NEK solo atiende a la segunda.

Por ello, diversos desarrollos han reformulado la CP-NEK para incorporar inercias, y se agrupan en dos familias:

- *Curvas híbridas con inercia de precios*: Gordon (1982) introduce retardos de la inflación en su "modelo del triángulo"[42], que combina, en la terminología posterior, la persistencia *extrínseca* (heredada de los determinantes de demanda y de oferta) y la *intrínseca* (procedente de la propia inflación). Por su parte, Galí y Gertler (1999) obtienen, a partir de Calvo con una fracción de empresas que siguen reglas sencillas, una CP-NEK híbrida con un retardo de la inflación junto a la inflación esperada y el coste marginal real. Por último, Christiano, Eichenbaum y Evans (2005) introducen la indexación generalizada de los precios no reoptimizados a la inflación pasada.

- *Inercia de información* (Mankiw y Reis, 2002): sustituye la rigidez de precios por la rigidez de información (*sticky information*), según la cual los agentes actualizan sus expectativas solo de forma intermitente. De ahí que, si los precios cambian con rapidez, únicamente una fracción de los agentes reaccione con información al día.

> Las curvas de Phillips híbridas, con el modelo del triángulo de Gordon, y las fuentes de la persistencia de la inflación se desarrollan en el tema A41, apartados 2.4.3 y 1.2.3, respectivamente.

3.4.2. Ausencia del sistema financiero (Cúrdia y Woodford, 2010)

Vasco Cúrdia y Michael Woodford reconocen como una deficiencia el no haber incorporado el sistema financiero al diseño de la política monetaria, y proponen un nuevo diseño de los modelos neokeynesianos que recoja las fricciones crediticias. En esta línea, los modelos con *acelerador financiero* (Bernanke, Gertler y Gilchrist, 1999) muestran cómo las imperfecciones del crédito amplifican las perturbaciones reales.

3.4.3. Desconsideración de la política fiscal (Solow, 2004)

Robert Solow (2004) señala que las discusiones sobre política fiscal habían quedado fuera del debate macroeconómico, dominado por el consenso (vigente hasta la Gran Recesión) de aplicar políticas fiscales neutrales que no interfiriesen con la política monetaria y evitasen la dominancia fiscal. Solow defiende que se subestimó el papel de la política fiscal como instrumento anticíclico, salvo ante grandes *shocks* que desembocasen en una trampa de liquidez, con los tipos nominales en el ZLB.

3.4.4. Recomendaciones metodológicas (Blanchard, 2008)

Olivier Blanchard (2008) aboga por cierta simplificación metodológica y por un mayor contacto de los modelos con los datos. En línea con Solow, observa que se ha prestado más atención a los estudios de política monetaria que a los de política fiscal. Como consecuencia, los debates posteriores a la crisis se han alejado en cierta medida de los planteamientos de la NEK.

[42] Se llama "del triángulo" porque la inflación depende de tres elementos: la inercia (la base del triángulo), la demanda (a través de la brecha de producción y el desempleo) y la oferta.

3.5. La nueva síntesis neoclásica

El recorrido de la escuela desemboca en la convergencia que Goodfriend y King (1997) bautizaron como **nueva síntesis neoclásica**. Woodford (2009) la describe como un consenso en tres planos:

- *Metodología*: modelos de equilibrio general dinámico y estocástico, con expectativas racionales y validez estructural frente a la crítica de Lucas.

- *Impulso del ciclo*: los *shocks* reales son una fuente importante de fluctuaciones, sin que ello excluya los nominales.

- *Política económica*: la política monetaria es efectiva, sobre todo para controlar la inflación.

Las críticas anteriores marcan la agenda posterior a la Gran Recesión, con fricciones financieras, heterogeneidad de agentes y un papel mayor de la política fiscal.

> Los principios del nuevo consenso y su génesis desde la nueva macroeconomía clásica se desarrollan en el tema A06, apartado 3.4.

Conclusiones

La nueva economía keynesiana nace como respuesta a la nueva macroeconomía clásica y acaba compartiendo con ella el método. Su primera generación microfundamentó dos tipos de rigideces. Las reales explican por qué el salario o el precio deseado apenas responde a la actividad, y las nominales, por qué no se ajusta el precio efectivo. La lección de Ball y Romer es que ambas se necesitan, porque sin rigidez nominal el dinero no tiene efectos reales y sin rigidez real la nominal es demasiado pequeña para importar. La segunda generación adoptó el marco del ciclo real y le añadió la competencia monopolística y la fijación escalonada de precios *à la* Calvo, de donde surge el modelo canónico de tres ecuaciones. En él la política monetaria tiene efectos reales, toma como referencia el *output* y el tipo de interés naturales y se guía por una regla que respeta el principio de Taylor, y la gestión de las expectativas forma parte de la política monetaria. La divina coincidencia se rompe con rigideces salariales reales o *shocks* de costes, y la política fiscal recupera eficacia en el límite inferior de los tipos.

Junto con la nueva macroeconomía clásica, la NEK forma la nueva síntesis neoclásica, que toma el método de la primera y las rigideces de la segunda (tema A06). Sus modelos de fijación de salarios y precios sustentan las teorías del desempleo de equilibrio y de la inflación, en las que la NAIRU resulta de la pugna por el reparto de la renta (tema A28). A su vez, el modelo de tres ecuaciones es el marco de referencia de la política monetaria. Explica por qué el dinero no es neutral a corto plazo (tema A37), qué regla minimiza la pérdida del banco central (tema A36), cómo interactúan la política monetaria y la fiscal (tema A40) y cómo se propagan las perturbaciones a lo largo del ciclo (tema A42).

Las críticas al modelo básico, como la escasa persistencia de la CP-NEK o la ausencia del sistema financiero, han marcado la agenda de la escuela sin desplazarla. En los modelos neokeynesianos con agentes heterogéneos (HANK) de Kaplan, Moll y Violante (2018), la política monetaria afecta al consumo sobre todo a través de efectos indirectos, vía la renta del trabajo, más que por la sustitución intertemporal. Con datos regionales de EE. UU., Hazell, Herreño, Nakamura y Steinsson (2022) estiman una curva de Phillips de pendiente pequeña y atribuyen la estabilidad de la inflación sobre todo al anclaje de las expectativas.

Anexo. Microfundamentación del modelo de la NEK 2G

Las tres ecuaciones del modelo canónico expuestas en el tema proceden de la optimización de hogares y empresas. Lo que sigue, de carácter opcional, esboza los pasos de esa derivación y numera (1)–(8) las ecuaciones a las que esta remite.

A.1. Agentes y equilibrio

Hogares: el hogar representativo elige consumo, oferta de trabajo y demanda de bonos.

- *Planteamiento*: maximiza su utilidad intertemporal esperada:

$$U = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \, u(c_t, l_t), \qquad \beta \in (0,1)$$

La utilidad instantánea es isoelástica (a partir de ahora se omiten los subíndices temporales):

$$u = \frac{c^{1-\theta}}{1-\theta} - \frac{l^{1+\varphi}}{1+\varphi}, \qquad \theta > 0$$

Esta función es continua y dos veces diferenciable, con $u'_c > 0$, $u''_c \leq 0$, $u'_l \leq 0$ y $u''_l \leq 0$. Sus variables y parámetros son:

- *Consumo*: es un agregador de Dixit-Stiglitz sobre un continuo de bienes diferenciados:

$$c = \left[\int_0^1 c(i)^{1-1/\sigma} \, di \right]^{\sigma/(\sigma-1)} \tag{1}$$

  donde $c(i)$ es la cantidad del bien $i \in [0,1]$ consumida en el periodo y σ la elasticidad de sustitución entre bienes diferenciados.

- l son las horas de trabajo ofrecidas por el consumidor.

- $\theta \equiv -\varepsilon_{u'_c c}$ es el coeficiente de aversión relativa al riesgo (elasticidad de la utilidad marginal del consumo respecto del consumo).

- $\varphi \equiv \varepsilon_{u'_l l}$ es la elasticidad de la desutilidad marginal del trabajo respecto del trabajo (inversa de la elasticidad de Frisch).

El consumidor está sujeto a su restricción presupuestaria, en la que precios y salarios se toman como dados:

$$P\,c + P_B\,b \leq b_{-1} + w\,l + Z$$

donde $P = \left[\int_0^1 p(i)^{1-\sigma}\, di\right]^{1/(1-\sigma)}$ es el índice de precios del consumo (dual del agregador de Dixit-Stiglitz), P_B el precio de los bonos, b un *stock* de bonos de riesgo nominal nulo y b_{-1} el heredado del periodo anterior, w el salario nominal, l las horas trabajadas y Z las transferencias netas del Gobierno y otras rentas. En suma, el gasto en consumo y bonos no debe superar a los ingresos. Se añade, además, una condición de *no-Ponzi* que impide endeudarse indefinidamente: $\lim_{T\to\infty} \mathbb{E}_t[\gamma_{t,T}\, b_T] \geq 0$, donde $\gamma_{t,T}$ es el factor de descuento estocástico entre t y T.

▪ *CPO*: con solución interior y la restricción agotada, derivando el lagrangiano respecto a c, l y b, se obtienen las dos condiciones de primer orden habituales, la oferta intratemporal de trabajo (2) y la ecuación de Euler de los bonos (3):

$$\underbrace{\frac{-\partial u/\partial l}{\partial u/\partial c}}_{\text{RMS}} = c^{\theta}\, l^{\varphi} = \underbrace{\frac{w}{P}}_{\text{salario real}} \tag{2}$$

$$P_B = \beta\, \mathbb{E}\left[\frac{\partial u/\partial c_{+1}}{\partial u/\partial c} \cdot \frac{P}{P_{+1}}\right] \tag{3}$$

▪ *CSO*: la utilidad es estrictamente cóncava en el consumo y en el trabajo (la desutilidad del trabajo es estrictamente convexa, con $\theta, \varphi > 0$) y la restricción presupuestaria es lineal. Por ello, con la condición de transversalidad (la de no-Ponzi satisfecha con igualdad), (2) y (3) son necesarias y suficientes.

Empresas: existe un continuo de empresas, cada una de las cuales produce un bien diferenciado (quedan indexadas igual que los bienes, $i \in [0, 1]$) y comparte la misma tecnología:

$$y_i = A\, l_i^{1-\alpha}$$

Cada empresa se enfrenta a la demanda de tipo Dixit-Stiglitz del hogar representativo, derivada del agregador de consumo (1):

$$c_i = \left[\frac{p(i)}{P}\right]^{-\sigma} c$$

▪ *Planteamiento, CPO y CSO*: el programa de fijación de precios *à la* Calvo, del que se obtiene la ecuación dinámica de los precios, se desarrolla en el apartado 1.4.1.

Sector público: en Galí (2008) no aparece el sector público, pero puede incluirse como una variable exógena.

▪ *Planteamiento*: como los bienes están diferenciados, el gasto público se agrega análogamente al consumo: $G = \left[\int_0^1 G(i)^{1-1/\sigma}\, di\right]^{\sigma/(\sigma-1)}$. A partir de aquí puede añadirse la restricción presupuestaria intertemporal del Gobierno y una regla de ajuste fiscal (p. ej. que recoja las desviaciones de impuestos y deuda respecto de sus valores naturales u objetivo).

Equilibrio: para obtener las tres ecuaciones se parte de las condiciones de vaciado en los mercados de bienes y de factores:

$$y = c + g \tag{4}$$

$$l = \int_0^1 l_i^d\, di \tag{5}$$

La condición (4) establece que la renta de la economía se destina íntegramente a sus usos posibles (aquí, consumo y gasto público). (5) establece que la oferta de trabajo del hogar representativo coincide con la demanda de trabajo agregada de las empresas que producen cada bien diferenciado.

Log-linealizando, se llega a una relación entre el *output* agregado, el empleo y la tecnología, que es la base para obtener la CP-NEK:

$$y = a + (1-\alpha)\, l \tag{6}$$

A.2. Obtención de la IS dinámica

Se parte de la ecuación de Euler del consumidor (3), particularizada con la utilidad isoelástica:

$$P_B = \beta\, \mathbb{E}\left[\frac{c_{+1}^{-\theta}}{c^{-\theta}} \cdot \frac{P}{P_{+1}}\right]$$

Log-linealizando en torno al estado estacionario[43], es decir, sustituyendo la expresión, originalmente no lineal, por una lineal mediante una aproximación de Taylor de primer orden en logaritmos, se obtiene la ecuación de Euler en consumo y, usando (4), en términos de renta:

$$c = \mathbb{E}[c_{+1}] - \frac{1}{\theta}(i - \mathbb{E}[\pi_{+1}] - \rho)$$

$$y = \mathbb{E}[y_{+1} - g_{+1}] - \frac{1}{\theta}(i - \mathbb{E}[\pi_{+1}] - \rho) + g$$

Expresada en desviaciones (brecha de producción) y definiendo el tipo de interés real natural $\bar{r}$, resulta la IS dinámica:

$$\tilde{y} = \mathbb{E}[\tilde{y}_{+1}] - \frac{1}{\theta}(i - \mathbb{E}[\pi_{+1}] - \bar{r})$$

A.3. Obtención de la CP-NEK

Partiendo de (6) se construyen dos relaciones que conectan el coste marginal de las empresas con el *output* agregado, para enlazar después *output*, desempleo e inflación. La ecuación (7) expresa el coste marginal real en función del salario real, la tecnología y el *output*, y la (8) relaciona el coste marginal medio con la producción agregada.

Para (7), el coste marginal real (en logaritmos) equivale al coste laboral marginal menos la productividad marginal del trabajo, que se obtiene de (6). Como $\mathrm{Pmg}_L = (1-\alpha) A\, l^{-\alpha}$, en logaritmos $\mathrm{pmg}_l = a - \alpha l + \ln(1-\alpha)$, de modo que:

$$\mathrm{cmg} = (w-p) - \mathrm{pmg}_l = (w-p) - a + \alpha l - \ln(1-\alpha) = (w-p) - \frac{1}{1-\alpha} a + \frac{\alpha}{1-\alpha} y - \ln(1-\alpha) \tag{7}$$

donde la última igualdad usa $l = (y-a)/(1-\alpha)$, despejado de (6).

Para (8), combinando el coste marginal real con las condiciones de equilibrio (oferta de trabajo y vaciado de mercados), la desviación del coste marginal medio respecto de su estado estacionario resulta proporcional a la brecha de producción:

$$\widehat{\mathrm{cmg}}_t = \eta\, \tilde{y}_t, \quad \eta = \theta + \frac{\varphi + \alpha}{1-\alpha} > 0 \tag{8}$$

donde el gorro denota desviaciones respecto del estado estacionario. El coste marginal es, por tanto, procíclico, y lo es más cuanto menos elástica es la oferta de trabajo (mayor φ) y más acusados son los rendimientos decrecientes (mayor α).

[43] Para una exposición de la técnica de log-linealización en torno al estado estacionario, habitual en los modelos del ciclo real (RBC) y DSGE, véase Galí (2008).

Sustituyendo (8) en la curva de Phillips en función del coste marginal, obtenida de la fijación de precios *à la* Calvo (apartado 1.4.1), resulta la curva de Phillips de la NEK en función de la brecha de producción:

$$\pi_t = \beta\,\mathbb{E}_t[\pi_{t+1}] + \kappa\,\widetilde{y}_t, \quad \kappa = \kappa_{\text{cm}}\,\eta, \qquad \kappa > 0,\ \beta \in (0,1)$$