1. Concepto, costes y efectos sobre el bienestar

El análisis empieza por definir y medir la inflación y sigue con sus costes, entre ellos los que impone su persistencia, y sus beneficios, de cuya comparación resulta la inflación óptima.

1.1. Concepto y medición

Definición: la **inflación** es el aumento sostenido del nivel general de precios. El carácter sostenido la distingue de una subida aislada, que altera el nivel de precios y solo transitoriamente su tasa de variación. Como fenómeno dinámico, exige considerar las expectativas.

Medición: hay tres medidas principales.

- El **IPC** es un índice de Laspeyres de una cesta representativa del consumo, nacional e importado, y es la referencia de los objetivos de inflación y de la indiciación de salarios y pensiones. El BCE usa el IPCA y la Reserva Federal, el deflactor del consumo privado.
- El **deflactor del PIB** es un índice de Paasche de lo producido en el país.
- La **inflación subyacente** excluye la energía y los alimentos no elaborados. Aproxima la tendencia de los precios sobre la que actúa la política monetaria.

Distinciones: por intensidad, es moderada, de un dígito, galopante, de dos o tres, o hiperinflación, por encima del 50 % mensual; por previsibilidad, esperada o no esperada, con costes distintos; por su causa, de demanda, de costes, estructural, inercial o monetaria. La relevante para la política es la sostenida y esperada, porque la no esperada no puede mantenerse con una inflación estable, pues los agentes acaban anticipándola.

1.2. Los costes de la inflación

Idea: los costes dependen de la magnitud, la volatilidad y la anticipación. Los de la inflación esperada subsisten aunque se anticipe, y los de la no esperada proceden de la sorpresa.

1.2.1. Costes microeconómicos

Conducta de los agentes: en una expansión, la inflación puede adelantar el consumo y la inversión y alimentar burbujas, porque los agentes anticipan sus compras al esperar precios más altos. Además, la *ilusión monetaria* lleva a confundir magnitudes nominales y reales, de modo que los agentes, al ignorar muchos efectos de la inflación, no optimizan adecuadamente.

Mercado de bienes: salvo que sea pura, la inflación hace que los precios informen peor de la escasez relativa, con costes de búsqueda e incertidumbre, y obliga a revisar precios, con **costes de menú**.

Mercado de trabajo y de factores: reduce el valor real de los salarios no ajustados y de la riqueza nominal, y, si se concentra en la energía o las materias primas, altera las cadenas de valor. Esa pérdida de poder de compra puede, además, enfriar el consumo.

Mercados financieros: afecta a las tres funciones del dinero. Como medio de pago se acepta peor, como unidad de cuenta pierde fiabilidad y como depósito de valor pierde poder de compra. Surgen así los **costes de suela de zapatos**, porque, al reducir el efectivo, los agentes acuden a otros activos y multiplican las conversiones en saldos líquidos, con sus costes de transacción. Su coste de bienestar (Bailey, 1956) es el área entre la demanda de saldos reales y el eje del tipo, de $i = 0$ al vigente, menos la transferencia $i\,m(i)$. Con la demanda de Cagan crece aproximadamente con el cuadrado del tipo nominal, ronda el 1 % de la renta al pasar de una inflación del 10 % a cero (Lucas, 2000) y desaparece con la regla de Friedman. Además, la inflación no anticipada redistribuye arbitrariamente de acreedores a deudores, y la alta y volátil añade primas de riesgo a los tipos nominales, que encarecen la financiación en términos reales.

1.2.2. Costes macroeconómicos

Tipo de interés real: con inflación esperada, $i = r + \pi^e$, y a largo plazo la inflación se traslada íntegramente a los tipos nominales, el **efecto Fisher**. La no esperada reduce el tipo real *ex post* y favorece a los deudores, mientras que a largo plazo el tipo real queda determinado por factores reales.

Sector exterior: con tipo fijo o traslación incompleta, una inflación superior a la de los socios resta competitividad-precio y deteriora la balanza comercial si se cumple Marshall-Lerner.

Crecimiento: la incertidumbre sobre los precios relativos y la rentabilidad acorta los horizontes y reduce la inversión. La evidencia asocia inflaciones elevadas y volátiles con menor crecimiento, aunque el efecto solo es apreciable por encima de las tasas moderadas.

1.2.3. Costes desde la política económica

Política monetaria: una inflación alta y volátil desancla las expectativas, con lo que la curva de Phillips de corto plazo se desplaza y la política pierde capacidad de estabilizar. Además, volver a la estabilidad cuesta producción y empleo, la *ratio de sacrificio*, que mide los puntos de producción perdidos por cada punto de inflación reducida.

Persistencia: la inflación tiende a permanecer cerca de su valor anterior, como la inercia física, y el coste de desinflar depende de sus fuentes (Fuhrer, 2006).

- La **persistencia intrínseca** procede de la propia inflación y hace las desinflaciones más largas y costosas y la respuesta a los *shocks* más persistente.
- La **persistencia heredada** procede de determinantes con dinámica propia, como el desempleo, la brecha o el coste marginal.

La distinción importa porque la heredada desaparece con sus causas, mientras que la intrínseca obliga a desinflar de forma gradual y costosa. Qué parte es intrínseca es el punto en que discrepan las teorías, y el criterio con el que la evidencia las juzga.

Política fiscal: dos distorsiones.

- El **efecto Tanzi-Olivera**: como se recauda con demora, el valor real de lo recaudado cae por debajo del gasto y el déficit se amplía.
- La **progresividad en frío**: con impuestos progresivos no indiciados, las subidas nominales llevan a tramos superiores, aunque el coste se debe al diseño del impuesto, no a la inflación.

En suma, la inflación esperada cuesta suela de zapatos, menú y distorsiones fiscales, pequeños a tasas moderadas; la no esperada, redistribución arbitraria, incertidumbre y pérdida de información de los precios. Por eso una inflación moderada y estable cuesta menos que una baja pero volátil. Ambos tipos de costes justifican que la estabilidad de precios sea el objetivo central de la política monetaria. La solución a muchos de estos costes pasa por la indiciación de salarios, tramos impositivos y deuda.

1.3. Beneficios e inflación óptima

Microeconómicos: con tasas bajas, la conducta **casi racional** de quien no incorpora la inflación a sus decisiones permite un desempleo sostenible menor que con inflación nula (Akerlof et al., 2000). El beneficio se agota, sin embargo, a tasas más altas, cuando ignorar la inflación sale caro y los agentes la incorporan a sus contratos. Con salarios nominales rígidos a la baja, la inflación reduce los salarios reales sin recortar los nominales y "engrasa" el mercado de trabajo, como sostenía Tobin. Además, grava indirectamente la economía sumergida, que estrecha las bases imponibles y se paga sobre todo en efectivo, de modo que la inflación alcanza una base que los impuestos ordinarios no tocan. Esto justifica una inflación óptima positiva, aunque la ganancia es pequeña.

Macroeconómicos: a corto plazo, si $\pi^e > i$, el tipo real es negativo y estimula la inversión. Por el **efecto Mundell-Tobin**, la inflación esperada desplaza la cartera hacia activos reales, eleva la acumulación de capital y reduce el tipo real. Por este canal puede tener efectos reales permanentes en un modelo con dinero, aunque pequeños. Y, por el mismo efecto que perjudica a los acreedores, reduce el valor real de las deudas de hogares, empresas y Estado, lo que facilita salir de las crisis de sobreendeudamiento y es la contrapartida de la deflación de deuda.

Política económica: una inflación positiva aleja los tipos nominales del límite inferior y da margen para bajarlos en las recesiones. Además, el *señoraje*, el ingreso real que el Estado obtiene por la emisión de moneda, permite financiar gasto sin subir impuestos y completa con los impuestos y la deuda sus vías de financiación.

Inflación óptima: la **regla de Friedman** (1969) sitúa el óptimo en un tipo nominal nulo, es decir, por la relación de Fisher, en una deflación igual al tipo real:

$$i = r + \pi^e = 0 \quad \Longleftrightarrow \quad \pi^e = -r$$

Solo así el coste de mantener dinero iguala su coste de producción, que es nulo. El consenso se aparta de ella por tres razones: el sesgo al alza de los índices de precios, la rigidez salarial a la baja y el margen frente al límite inferior de los tipos. De ahí procede el objetivo del 2 %, una inflación reducida, controlada y anticipable.

En suma, la inflación no es neutral y su análisis fija el objetivo de la política monetaria, aunque los costes a tasas moderadas parecen reducidos (Lucas, 2000) y el nivel óptimo, del 2 %, el 3 % o el 4 %, sigue en debate.

2. Las causas de la inflación: escuelas y curvas de Phillips

Las causas se ordenan en cinco tipos: de demanda, de costes, estructural o de puja distributiva, inercial o de expectativas, y monetaria. Las escuelas difieren en cuál destacan y en cómo la fundamentan.

2.1. Teoría clásica y teoría keynesiana

2.1.1. La teoría cuantitativa del dinero

Origen: nace en la Escuela de Salamanca, con Martín de Azpilcueta (1556), que atribuyó la subida de precios a la plata americana, antes que Jean Bodin (1568). Su formulación moderna se debe a Fisher (1911) y a la escuela de Cambridge, que dieron forma algebraica a la relación causal de Hume (1752).

Idea: a largo plazo, precios y dinero se mueven en la misma proporción, de modo que la inflación es siempre un fenómeno monetario.

Ecuación de cambio (Newcomb, 1885; Fisher, 1911): es una identidad, $M \cdot V_T = P \cdot T$, que Fisher convierte en relación causal al suponer

estables la velocidad-transacción, fijada por la tecnología de pagos, y las transacciones, fijadas por factores reales: $P = (V_0/T_0)\,M$. El dinero es un velo: se cumplen la **neutralidad monetaria** y la **dicotomía clásica**.

Ecuación de Cambridge (Marshall, Pigou): $M = k \cdot P \cdot Y$, con $k = 1/V_Y$. La conclusión es la misma, pero se lee como una demanda de dinero, y k, una decisión de los agentes, puede depender del tipo de interés y de las expectativas, de lo que partirá Keynes. En ambas, M es la cantidad de dinero, P el nivel general de precios, T el volumen de transacciones e Y la renta real.

De cualquiera de las dos versiones se sigue que, con dinero exógeno, sus cambios se trasladan proporcionalmente a los precios a largo plazo. Esos cambios solo afectan al nivel de precios, que es la neutralidad, y las variables reales se determinan con independencia de las monetarias, que es la dicotomía.

Fija la conclusión de largo plazo que nadie ha desmentido: la inflación sostenida es un fenómeno monetario. Pero, al comparar equilibrios estáticos, explica el nivel de precios, o un salto aislado, mejor que su subida sostenida, que exigiría una expansión monetaria mantenida que la propia neutralidad deja sin justificación.

2.1.2. La teoría keynesiana: demanda y costes

Idea: la Gran Depresión puso en duda el paradigma neoclásico. Con rigideces de precios y salarios, la inflación no es solo monetaria. Más dinero puede tener efectos reales, porque con salarios nominales rígidos abarata el salario real, y la inflación puede nacer de la demanda o de los costes con independencia del dinero. Distingue dos orígenes:

- La **inflación de demanda**, formalizada con la *brecha inflacionista* por Keynes (1940) y Smithies (1942), aparece en el pleno empleo o cerca de él, cuando la oferta no puede seguir a la demanda. Los consumidores compiten entonces por los productos y presionan al alza los precios.
- La **inflación de costes** aparece aun con desempleo elevado, por la fijación de *mark-ups* con poder de mercado, como cuando suben los precios del petróleo y, con ellos, los costes de todas las empresas que usan el combustible.

Sigue siendo útil para el diagnóstico, porque una misma inflación puede requerir respuestas distintas según su origen, pero también compara equilibrios estáticos y no explica la dinámica. Una subida continuada de los precios exigiría que el exceso de demanda, o la escasez de oferta, durase también.

2.2. Síntesis neoclásica y monetarismo

2.2.1. Síntesis: la curva de Phillips

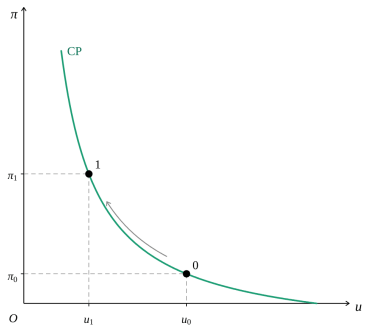
Origen: Phillips (1958) halló una relación inversa entre el crecimiento de los salarios nominales y el desempleo en el Reino Unido, Lipsey (1960) la fundamentó y Samuelson y Solow (1960) la convirtieron en una relación entre inflación y desempleo.

Idea: la inflación nace de la presión de la demanda sobre el mercado de trabajo. La relación es estable y ofrece un menú a la política económica. Por la ley de Okun:

$$\pi_t = -\kappa_u(u_t - \bar{u}) \qquad \Longleftrightarrow \qquad \pi_t = \kappa\,(y_t - \bar{y}), \qquad \kappa_u > 0,\ \kappa > 0$$

con $\kappa_u, \kappa > 0$, de modo que la inflación aparece cuando la producción supera a su potencial, se anula cuando lo iguala y se vuelve negativa cuando queda por debajo.

Desarrollo gráfico: en el plano desempleo-inflación la curva es decreciente, y una expansión monetaria o fiscal lleva la economía a lo largo de ella del punto 0 al 1, con $u_1 < u_0$ y $\pi_1 > \pi_0$.



Curva de Phillips original: una expansión de la demanda lleva la economía de 0 a 1.

Es una teoría de demanda: una expansión sostenida mantiene la producción por encima del potencial a cambio de una inflación más alta pero estable. Se apoyaba, sin embargo, en que los trabajadores no incorporaban la inflación a sus demandas, y la estanflación de los setenta mostró que la curva se había desplazado.

2.2.2. Monetarismo: la curva ampliada

Origen: Friedman recuperó la teoría cuantitativa como teoría de la demanda de dinero (1956) e introdujo las expectativas en la curva de Phillips (1968), a la vez que Phelps (1967).

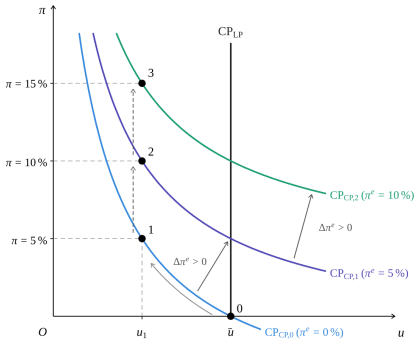
Reformulación cuantitativa: la demanda de dinero depende de la renta permanente y es estable. En tasas, $\pi = g_M + g_V - g_Y$. Con velocidad estable, la inflación es el exceso del crecimiento monetario sobre el real, y es "siempre y en todas partes un fenómeno monetario" (Friedman, 1963). Las causas de demanda o de costes solo la sostienen si el dinero las acomoda. Friedman lee la ecuación de Cambridge como una demanda de dinero, $M/P = k(\cdot)\,Y$, que depende también de los rendimientos de los activos alternativos y es más estable que la función de consumo keynesiana.

Curva ampliada por las expectativas: si los trabajadores negocian en términos reales, la curva se desplaza con la inflación esperada, con $\bar{u}$ la tasa natural de paro:

$$\pi_t = \pi_t^e - \kappa_u(u_t - \bar{u}), \qquad \kappa_u > 0$$

- El **componente inercial**: con expectativas adaptativas, la inflación esperada es la pasada y se autoalimenta.
- El **componente de demanda**: solo la producción por encima del potencial presiona los precios. La curva original es el caso $\pi_t^e = 0$.
- La **aceleración**: mantener el desempleo por debajo de la tasa natural exige una inflación creciente, sostenida por un crecimiento monetario cada vez mayor, porque cada periodo la esperada alcanza a la observada.

Dinámica: a corto plazo, cada curva corresponde a una inflación esperada. A largo plazo, la curva es vertical en $\bar{u}$. Si la autoridad fija $u_1 < \bar{u}$, la inflación pasa al 5 %. Al revisarse las expectativas, la curva sube y hace falta un 10 %, y luego un 15 %. Cuando deja de acelerar, la economía vuelve a $\bar{u}$ con la inflación alcanzada. Las curvas de corto plazo son decrecientes porque, dada la inflación esperada, solo una inflación superior a ella lleva el desempleo por debajo de la tasa natural.



Hipótesis aceleracionista: mantener $u_1 < \bar{u}$ exige una inflación creciente (puntos 1, 2 y 3).

Es el punto de partida de toda la teoría moderna de la inflación, pero las expectativas adaptativas suponen errores sistemáticos y la tasa natural queda sin explicar, lo que aportará el modelo de Layard y Nickell.

2.3. Nueva macroeconomía clásica

2.3.1. Información imperfecta (Lucas, 1972)

Idea: con expectativas racionales, los agentes usan toda la información y sus errores no siguen un patrón sistemático. Si aun así existe un *trade-off* a corto plazo, es porque su información sobre los precios agregados es incompleta, un problema de extracción de señales.

Razonamiento: los productores no distinguen si una subida de su precio es un *shock* específico, ante el que convendría producir más, o general, ante el que no. Atribuyen al componente específico una fracción, el **coeficiente de extracción de señales**:

$$\vartheta = \frac{\sigma_i^2}{\sigma_i^2 + \sigma_p^2}$$

donde σ_i^2 es la varianza de los precios relativos y σ_p^2 la del nivel general, y cuanto mayor es la primera, más verosímil es que el *shock* sea específico.

Con inflación muy variable, ϑ es cercana a cero y los productores no ajustan la producción, mientras que con precios estables las sorpresas tienen efectos reales transitorios. La oferta agregada es $y_t = \bar{y}_t + \zeta\,(p_t - p_t^e) + \varepsilon_t^s$: solo las sorpresas tienen efectos reales, la política sistemática es inefectiva (Sargent y Wallace, 1975) y la inflación anticipada no compra actividad. Con $\zeta > 0$ creciente en ϑ, el efecto de las sorpresas es menor cuanto más variable es la inflación, y una autoridad que intente sorprender de forma sistemática deja de sorprender.

Aspectos positivos: sitúa las expectativas y la credibilidad en el centro, con una implicación contrastable entre países.

Limitaciones: la confusión es poco plausible con índices mensuales, no explica la persistencia y Mishkin (1982) halla efectos reales de la política anticipada.

2.3.2. Inconsistencia dinámica (Barro y Gordon, 1983a)

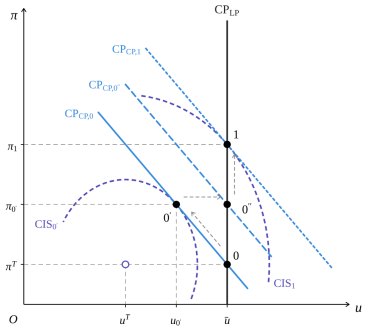
Idea: Kydland y Prescott (1977) mostraron que una política anunciada es **dinámicamente inconsistente** si, una vez que los agentes han actuado, a la autoridad le conviene incumplirla. Barro y Gordon (1983a) lo aplican a la inflación. La falta de credibilidad genera por sí sola un **sesgo inflacionista**, sin ganancia de empleo.

Supuestos: los agentes tienen expectativas racionales y fijan sobre ellas sus contratos. El Gobierno minimiza una pérdida social, con un objetivo de inflación, π^T, y otro de desempleo inferior a la tasa natural, $u^T < \bar{u}$. Las expectativas son $\pi_t^e = \mathbb{E}_{t-1}[\pi_t]$, y las curvas de indiferencia social son elipses en torno a (u^T, π^T), tanto mejores cuanto más cerca de ese objetivo.

Planteamiento: en el punto 0, la autoridad anuncia que controlará la inflación y los agentes firman sus contratos. Con salarios fijados, le conviene sorprender con más inflación para bajar el desempleo, punto 0 . Los agentes revisan sus expectativas, la curva de corto plazo se desplaza, punto 0 , y, tras varios periodos, se llega al punto 1, un equilibrio de Nash en el que la curva de corto plazo es tangente a una curva de indiferencia social.

Equilibrio: como los agentes prevén la secuencia, reclaman desde el principio el salario del punto 1: mismo desempleo, el natural, y más inflación. El anuncio inicial no es un equilibrio perfecto en subjuegos. La inflación puede nacer, así, de que el público no se fíe de los anuncios de política económica.

Desarrollo gráfico: en el plano desempleo-inflación, la curva de largo plazo es vertical en $\bar{u}$ y las de corto plazo, una para cada inflación esperada, tienen pendiente negativa. En el punto 1, la $\text{CP}_{\text{CP},1}$ es tangente a la CIS_1 sobre la de largo plazo, con el desempleo en $\bar{u}$ y una inflación π_1 superior al objetivo π^T del punto 0.



Barro y Gordon: del compromiso (0) al equilibrio discrecional (1), con sesgo inflacionista y sin ganancia de empleo.

Estática comparativa: el sesgo es mayor cuanto más baja es la meta de desempleo respecto a la natural, cuanto más peso da el Gobierno al desempleo y cuanto más reduce el paro una sorpresa dada. En todos los casos el desempleo de equilibrio sigue siendo la tasa natural, y los parámetros solo alteran la inflación.

Extensiones: las soluciones al sesgo son la independencia del banco central, las reglas transparentes, el **banquero conservador** (Rogoff, 1985) y los contratos de desempeño (Walsh, 1995). El banquero conservador tiene más aversión a la inflación que la sociedad, y los contratos vinculan la remuneración del gobernador al cumplimiento del objetivo de inflación.

Aspectos positivos: fundamenta el diseño institucional moderno de la política monetaria.

Limitaciones: supone $u^T < \bar{u}$ y que se conocen las preferencias del Gobierno. Con incertidumbre surgen la reputación y la señalización (Backus y Driffill, 1985).

2.4. Nueva economía keynesiana

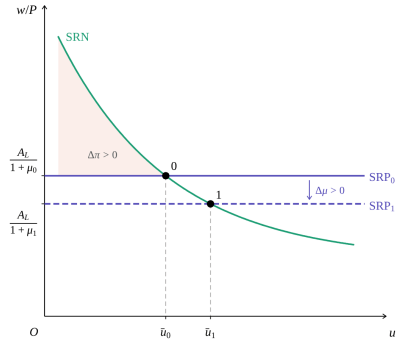
2.4.1. Puja distributiva (Layard y Nickell, 1985)

Idea: la inflación resulta de una **puja distributiva**. Los trabajadores negocian un salario real que crece con su poder y las empresas fijan un margen sobre el coste. Cuando las aspiraciones son incompatibles, porque el desempleo es bajo, la puja se resuelve en inflación. La tasa a la que son compatibles es la **NAIRU**.

Razonamiento: se obtiene la misma curva ampliada, pero $\bar{u}$ depende de parámetros estructurales: crece con el margen y con las prestaciones, el salario mínimo o las cotizaciones, y cae con la productividad. El salario negociado, $w/P^e = f(u, z)$, decrece con el desempleo, y el pagado, $A_L/(1+\mu)$, es horizontal. Por debajo de la NAIRU, el salario negociado supera al pagado y la inflación crece, mientras que por encima el negociado queda por debajo y la inflación se reduce. La NAIRU iguala ambos salarios, con A_L la productividad del trabajo y μ el margen:

$$f(\bar{u}, z) = \frac{A_L}{1+\mu}$$

Solo las políticas que actúan sobre la competencia en los mercados de bienes y de trabajo la reducen de forma permanente.



Puja distributiva: un mayor margen ($\Delta\mu > 0$) eleva la NAIRU de $\bar{u}_0$ a $\bar{u}_1$.

Explica inflaciones distintas con la misma política monetaria y fundamenta las políticas de rentas y las reformas estructurales, pero la NAIRU no es observable y el modelo no explica la dinámica del ajuste.

2.4.2. Fijación escalonada de precios (Calvo, 1983)

Origen: Calvo propuso la fijación escalonada como alternativa a los contratos de Fischer (1977) y Taylor (1980), y es hoy el estándar de los modelos DSGE. En él la revisión de precios es aleatoria y no depende del tiempo transcurrido desde la última, a diferencia de los contratos de duración fija, y es analíticamente más manejable.

Idea: las empresas tienen poder de mercado, pero cada periodo solo una fracción $1-\xi$ revisa su precio, de forma aleatoria, y lo fija con un margen sobre los costes marginales presentes y esperados. Agregando y loglinealizando, la **curva de Phillips neokeynesiana** es:

$$\pi_t = \beta \mathbb{E}_t[\pi_{t+1}] + \kappa\, \tilde{y}_t, \qquad \kappa > 0,\ \beta \in (0,1)$$

donde $\tilde{y}_t$ es la brecha de producción. En la calibración típica (Galí, 2008), $\xi = 2/3$, con precios fijos durante unos tres trimestres, y $\beta = 0{,}99$.

Con κ creciente en la frecuencia de ajuste, la inflación de hoy depende de la esperada para mañana y del exceso de demanda actual. Es puramente prospectiva: la inflación resume la senda esperada de brechas futuras, no hay inercia y la desinflación creíble no cuesta. Es la oferta de los modelos de los bancos centrales y hace de la credibilidad y la comunicación los instrumentos centrales, pero no reproduce la persistencia observada.

2.4.3. Curvas de Phillips híbridas

Críticas: la CP-NEK predice un ajuste inmediato, cuando la inflación cae con retraso tras una contracción (Mankiw, 2001), no explica la persistencia y las empresas apenas se informan sobre la inflación (Coibion et al., 2018). Por ello se considera que la inflación pasada predice mejor la futura que la esperada.

Desarrollos: recuperan la inercia del triángulo de Gordon.

- **Gordon** (1982): su modelo del triángulo formaliza la persistencia con tres vértices, la inercia, la demanda y la oferta, más un componente estocástico. Los determinantes son la inflación de los n periodos anteriores, ponderada por a_i, que recoge la persistencia intrínseca, como la hipótesis aceleracionista; el diferencial entre el desempleo y la NAIRU, $\bar{u}$, ponderado por b, y los factores de oferta, s_t, ponderados por c:

$$\pi_t = \sum_{i=1}^{n} a_i \pi_{t-i} - b(u_t - \bar{u}) + c \cdot s_t + \varepsilon_t$$

- **Galí y Gertler** (1999): reglas de ajuste sencillas, un retardo de la inflación y el coste marginal real, aproximado por la participación del trabajo en la renta, en lugar de la brecha. Esta especificación es la más usada.
- **Christiano, Eichenbaum y Evans** (2005): la indiciación a la inflación pasada de quienes no optimizan genera persistencia endógena.
- **Mankiw y Reis** (2002): rigidez de la información, no de los precios. Como informarse es costoso, cada periodo solo una parte de las empresas actualiza su información y el resto fija sus precios con información atrasada, lo que genera un ajuste gradual.

La curva híbrida es $\pi_t = \gamma_b \pi_{t-1} + \gamma_f \mathbb{E}_t[\pi_{t+1}] + \kappa \tilde{y}_t$. La desinflación tiene un coste, la ratio de sacrificio, mayor cuanto más pesa la inercia y menos credibilidad hay. El triángulo explica la estanflación y sirve para estimar la NAIRU y la ratio de sacrificio. Su límite es que la inercia descansa en retardos sin fundamento microeconómico, en la indiciación o en rigideces de información postuladas, expuestas a la crítica de Lucas. Además,

el peso estimado de la inercia frente a las expectativas varía con la especificación y con la medida del coste marginal.

En suma, las teorías convergen en $\pi_t = \pi_t^e - \kappa_u(u_t - \bar{u})$, coinciden a largo plazo con la teoría cuantitativa y difieren en cómo se forman las expectativas, de qué depende la tasa natural y de dónde vienen los efectos reales. En la NMC, la inflación anticipada no tiene efectos reales: Lucas explica los efectos de las sorpresas, y Barro y Gordon, la inflación por falta de credibilidad, que resuelven los bancos centrales independientes y los objetivos explícitos, cuyo éxito explica la baja inflación de 1990 a 2020. Ambas, con precios flexibles, predicen desinflaciones creíbles sin coste. Ninguna explica bien la persistencia sin supuestos *ad hoc*, ni la inflación importada y de oferta de los episodios recientes, y todas tratan la economía como cerrada.

3. Hiperinflación y deflación

En ambos extremos el dinero deja de cumplir sus funciones. En la hiperinflación pierde la de depósito de valor y acaba perdiendo la de medio de pago, y en la deflación su valor crece y los agentes lo atesoran. Ambos son fenómenos monetarios en su mecánica y fiscales o financieros en su origen.

3.1. La hiperinflación

3.1.1. Concepto y origen

Definición: según Cagan (1956), la **hiperinflación** supera el 50 % mensual, lo que multiplica los precios por más de cien en un año. Es rara y se asocia a guerras y colapsos del Estado: Alemania en 1923, Hungría en 1946, Zimbabue en 2008 o Venezuela entre 2017 y 2020.

Origen: es fiscal. Un Estado sin impuestos ni deuda recurre al señoraje. La inflación reduce la demanda de dinero, lo que obliga a emitir más para el mismo ingreso real, y el proceso se acelera con las expectativas. La hiperinflación es la forma extrema de la inflación monetaria, con la particularidad de que la aceleran las expectativas y la demanda de dinero, y no solo la oferta.

3.1.2. El modelo de Cagan (1956)

Supuestos: las variables reales son constantes, el Gobierno fija g_M para financiar el déficit y las expectativas son adaptativas.

Planteamiento: la demanda de saldos reales depende solo de la inflación esperada, que se revisa en proporción al error cometido:

$$m_t - p_t = c - a\pi_t^e$$
$$\dot{\pi}^e = \lambda\,(\pi - \pi^e)$$

Aquí $c > 0$ recoge la demanda de dinero por factores reales, $a > 0$ su sensibilidad a la inflación esperada y $\lambda > 0$ la velocidad de aprendizaje.

Equilibrio: la demanda de dinero da el nivel de precios, y su versión en tasas, $g_M - \pi = -a\,\dot{\pi}^e$, con la HEA, la inflación observada en función de la esperada:

$$p_t = m_t - c + a\pi_t^e$$
$$\pi = \frac{g_M - a\lambda\,\pi^e}{1 - a\lambda} \qquad \pi^e = \pi = g_M$$

En el estado estacionario, la inflación observada, la esperada y el crecimiento monetario coinciden.

Desarrollo gráfico: en el plano inflación esperada-inflación observada, la recta $\pi(\pi^e)$ tiene pendiente $-a\lambda/(1-a\lambda)$, negativa si $a\lambda < 1$ y positiva y mayor que uno si $a\lambda > 1$. Corta a la bisectriz, $\pi = \pi^e$, en el equilibrio, el punto 0, con $\pi^e = \pi = g_M$.

Estática comparativa: el nivel de precios es menor cuanto mayor es c y crece con la cantidad de dinero, y la inflación de equilibrio, con g_M. El modelo fue un poderoso argumento a favor de controlar los agregados monetarios. Cuando la inflación se acelera, los precios suben más cuanto mayores son a y λ, que intensifican la retroalimentación.

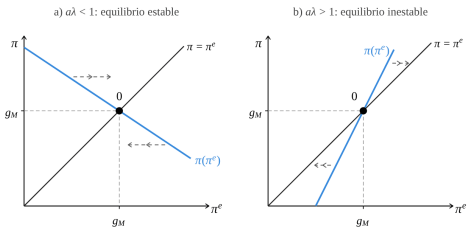
Dinámica: la inflación esperada evoluciona según $\dot{\pi}^e = \lambda(g_M - \pi^e)/(1 - a\lambda)$. Si $a\lambda < 1$, el equilibrio es estable y la hiperinflación solo se explica por una emisión excesiva y creciente. Si $a\lambda > 1$, un salto de la inflación eleva la expectativa, reduce la demanda de saldos y el exceso de dinero genera aún más inflación, de modo que la hiperinflación es una **profecía autocumplida**, sin que el dinero crezca de forma continua.

Con expectativas racionales, son las expectativas sobre el nivel de precios futuro las que lo determinan, y un Gobierno sin credibilidad perpetúa la inflación también por el lado de la demanda de dinero.

Señoraje: es $\sigma = g_M \cdot (M/P)$, el tipo por la base, y, en estado estacionario, $\sigma = g_M\, e^{c - a g_M}$. Aumentar g_M eleva el tipo pero reduce la base, porque los agentes huyen del dinero. La condición de primer orden da el máximo:

$$e^{c - a g_M}(1 - a g_M) = 0 \quad \Longrightarrow \quad g_M^* = \frac{1}{a}, \qquad \sigma^* = \frac{e^{c-1}}{a}$$

que crece con c y decrece con a. El máximo es la cima de una curva de Laffer, y las hiperinflaciones aparecen cuando el Gobierno quiere más señoraje del que permite la demanda de dinero.



Estabilidad en el modelo de Cagan: converge si $a\lambda < 1$ (a) y se autoalimenta si $a\lambda > 1$ (b).

El modelo formaliza por primera vez la hiperinflación y trata la inflación como un impuesto, pero supone una economía cerrada, y el máximo de señoraje no es el óptimo social. Con tipo de cambio fijo, monetizar el déficit haría perder reservas y acabaría forzando el abandono de la paridad. Tampoco endeudarse en lugar de emitir dinero evita la inflación, según la "aritmética monetarista desagradable" de Sargent y Wallace (1981).

3.1.3. Estabilización (Sargent, 1982)

Idea: las hiperinflaciones de Austria, Hungría, Polonia y Alemania terminaron de golpe, con un coste menor del que predecía la ratio de sacrificio, gracias a un **cambio de régimen** creíble que combinó:

1. Un ajuste fiscal que eliminó el déficit, la causa de la inflación. Como esta procedía de las expectativas de financiación monetaria del déficit, solo un cambio institucional visible podía cambiarlas.
2. Un banco central independiente que no podía financiar al Tesoro, lo que hacía creíble el ajuste.
3. Un ancla nominal, a menudo una nueva moneda o un tipo de cambio fijo, que daba a los agentes una referencia para fijar precios.

Los programas heterodoxos de los ochenta añadieron controles de precios y salarios, con éxito solo con ajuste fiscal. La credibilidad, no la contracción gradual, es el instrumento, aunque la conclusión se traslada peor a inflaciones moderadas, en las que los contratos y la indiciación hacen la desinflación más gradual y costosa.

3.2. La deflación

Definición: la **deflación** es la caída sostenida del nivel general de precios. Es benigna si procede de *shocks* de oferta positivos y recesiva si procede de *shocks* de demanda negativos. No es simétrica a la inflación por tres asimetrías: las deudas nominales, la rigidez salarial a la baja y el límite inferior de los tipos. La benigna acompaña a una mayor productividad o a importaciones más baratas, y la recesiva, a una crisis financiera o a una contracción del crédito.

Costes: Fisher (1933) explicó la Gran Depresión por la **deflación de deuda**: la caída de precios eleva la deuda real, los deudores venden activos, los precios caen más y, cuanto más pagan, más deben. La rigidez salarial eleva el salario real y el desempleo. Y en la **trampa de liquidez** (Keynes, 1936), con $i = 0$, la deflación esperada eleva el tipo real, $r = i - \pi^e$, cuando el banco central ya no puede bajar los tipos. La deflación recesiva es el escenario en que la política monetaria convencional es inoperante, y por eso los bancos centrales la temen más que a una inflación moderada. Japón desde los noventa es el caso de referencia, y la "japonización", el riesgo de una deflación prolongada con tipos en cero y crecimiento débil. La deflación de deuda es, además, precursora del acelerador financiero, aunque su efecto agregado depende de que los deudores gasten más que los acreedores. La experiencia reciente muestra, con todo, que el límite inferior efectivo está algo por debajo de cero y que las medidas no convencionales conservan cierta capacidad de actuación.

Política económica: la política monetaria recurre al *forward guidance*, a la expansión cuantitativa, a los tipos negativos y a los objetivos de nivel de precios o de PIB nominal. La fiscal recupera eficacia, porque con tipos en cero no hay expulsión. La experiencia de Japón, de la zona del euro en 2014-2016 y de EE. UU. tras 2008 muestra que las medidas no convencionales evitan la espiral, pero no siempre devuelven la inflación al objetivo, y que lo eficaz ha sido combinarlas con la política fiscal. La deflación es el argumento más fuerte a favor de la coordinación. Los objetivos de nivel de precios o de PIB nominal sirven porque comprometen al banco central a compensar la deflación pasada con inflación futura.

En suma, la hiperinflación muestra que la inflación es, en última instancia, fiscal y de expectativas, y la deflación, que la caída de precios es más peligrosa que su subida. Ambas justifican un objetivo positivo y creíble.